STOCHASTIC PROCESSES

Selected Papers of Hiroshi Tanaka

Edited by

Makoto Maejima & Tokuzo Shiga

World Scientific

STOCHASTIC PROCESSES

Selected Papers of Hiroshi Tanaka

STOCHASTIC PROCESSES

Selected Papers of Hiroshi Tanaka

Edited by

Makoto Maejima
Keio University, Japan

Tokuzo Shiga
Tokyo Institute of Technology, Japan

World Scientific
New Jersey • London • Singapore • Hong Kong

Published by

World Scientific Publishing Co. Pte. Ltd.
P O Box 128, Farrer Road, Singapore 912805
USA office: Suite 1B, 1060 Main Street, River Edge, NJ 07661
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

STOCHASTIC PROCESSES
Selected Papers of Hiroshi Tanaka

ISBN 981-02-4591-2

Printed in Singapore by Uto-Print

Hiroshi Tanaka

Preface

Hiroshi Tanaka is known for his outstanding contributions to the theory of stochastic processes. Since his first paper on diffusion processes in 1957, he has obtained brilliant mathematical results including construction of diffusion processes with continuous coefficients, Tanaka formula for Brownian motion, Skorohod equation on a convex domain, Boltzmann equations, diffusion processes in random media and so on.

On the occasion of his retirement from academic positions, we planned to collect his important contributions to the theory of stochastic processes in one volume, making his works available in a unified form to the mathematical community.

Besides his selected papers in this volume we asked Professors Henry McKean, Marc Yor, Shinzo Watanabe and Tanaka himself to write some essays on Tanaka's mathematics and personality.

Among Tanaka's works the most popular one would be the celebrated Tanaka formula for Brownian local time, which was never published. The story around the birth of Tanaka formula can be found in the essays of Marc Yor and Shinzo Watanabe.

As editors of this volume, we are grateful to World Scientific Publishing Co. Pte. Ltd. for the production of the volume. Our special thanks go to Professors Henry McKean, Marc Yor and Shinzo Watanabe for their writing interesting essays, to Professors Nobuyuki Ikeda and Yuji Ito for their valuable suggestions in the process of planning this volume, to Professors Yozo Tamura, Hideki Tanemura and Yuki Suzuki for their assistance throughout its preparation.

Makoto Maejima
Tokuzo Shiga

Contents

[†]Numbers in brackets refer to the Bibliography of Hiroshi Tanaka (see pages 425–428)

H. Tanaka: An Appreciation

by Henry McKean

It is a pleasant thing for me to write this little note to accompany the selected papers of H. Tanaka collected in this volume.

When I was young and fresh out of school, I visited Kyôto for a year (1957/58) at the invitation of K. Itô. My great-grandfather (whom I never knew) was an amateur of Japanese art, generally, and of Okyô in particular. Part of his collection came down to my family, so even when I was quite little, I was fascinated by Japan and things Japanese. The double chance to see Japan and to work with Itô was an irresistible piece of luck and I was pretty well prepared for that.

What I was not prepared for, but happily surprised by, was the enthusiastic crowd of young Japanese probabilists under Itô's wing who came faithfully, once a week, to hear the new developments in Brownian motion and diffusion stemming from the ideas of W. Feller and E.B. Dynkin — "sowing the seeds of diffusion in the mathematical fields of Japan" as Itô put it. H. Tanaka stood out in this company for the quickness of his understanding, for his mathematical taste, and for his quiet but strong personality. That was our first acquaintance. Later he visited me at MIT in Cambridge; in between, we collaborated on additive functions of the Brownian path ([5] †, 1961).

Tanaka's mathematical production is marked by the elegant use of common ideas employed in novel ways and with uncommon skill. No better example can be found than his 1963 paper ([7]) in which Itô's lemma is used with a bit of extra audacity to prove a wonderful formula, simple but highly effective, for the Brownian local time. I would like particularly to cite the deep series of papers ([15], [16], [19], [27]) on Boltzmann's equation for the Maxwellian gas and the Markov process that underlies it, a subject dear to my heart. This had been initiated by Kaç and studied somewhat by me, but it was Tanaka who really understood it correctly, and everything of Tanaka's shows like elegance, simplicity, depth.

This note is to express to Tanaka my admiration and friendship and to wish him health and happy productive years still to come.

---

†Numbers in brackets refer to the Bibliography of Hiroshi Tanaka (see pages 425–428)

From Local Times to Random Environments...

by Marc Yor

It is both a pleasure and an honor to have been asked to write a few words about the works of Professor H. Tanaka.

At least two aspects of Professor Tanaka's works have been of much interest to me and, more generally to many probabilists in France, for quite some time :

a) *The celebrated Tanaka formula, expressing the local time at* 0 *of linear Brownian motion as the last term of an extended Itô formula.*

Not only does this formula extend to Brownian local time taken at every level $a(\in \mathbb{R})$, which makes it a very powerful tool to deduce the Markovian properties of these local times with respect to a, i.e. the Ray-Knight theorems ([1], [2]), but as shown by P.A. Meyer in his "Cours sur les intégrales stochastiques" (Lecture Notes in Math., 511, Séminaire de Probabilités X, 1976), it extends easily to the framework of continuous semi-martingale local times.

The works of P. Mc Gill ([3]) and T. Jeulin ([4]) for instance, and later those of L.C.G. Rogers and J. Walsh ([5], [6]), concerning the spatially indexed filtration of Brownian excursions, give ample evidence of the power of Tanaka formula.

It is an interesting fact that Tanaka formula appears as a "personal communication" in H.P. McKean's "Stochastic Integrals" (1969) whilst more general discussions describing every additive functional of (possibly multidimensional) Brownian motion appeared earlier, e.g., in Professor Tanaka's paper (Zeitschrift für Wahr.,1963), and a number of papers by probabilists in the then Soviet Union.

There, I see some similarity with the role Problem 1 on p. 72 of Itô-McKean's book (1965), which establishes the existence of principal values of Brownian local times, has taken in the later development of these principal values (see, e.g. the discussion by T. Yamada [7], p. 414).

In the study of intersections of 2 - or 3 - dimensional Brownian motion, variants of Tanaka formula - the Tanaka-Rosen formulae - have also played some part in understanding Varadhan type renormalizations (see, e.g., [8]).

b) *The study of many properties - most of them asymptotic - of diffusions in random, especially white noise, environments.*

There are at least sixteen papers by Professor Tanaka, which have been devoted to this topic ; through his perseverance, many facts about these diffusions have been exposed in great clarity, in particular via the study of the so called "valleys".

For quite some time, I wished to establish some relationship between the $(\log t)^2$ normalization in the asymptotics of these diffusions, and the following radial version of Spitzer's theorem for the windings of two dimensional Brownian motion $(Z_t)_{t\geq 0}$:

$$\frac{4}{(\log t)^2}\int_0^t \frac{ds}{|Z_s|^2} \xrightarrow[t\to\infty]{(law)} T$$

where T denotes a stable $\left(\frac{1}{2}\right)$ variable.

More generally, I thought some link might exist between the one-dimensional diffusions in random environment (: RED) and some two-dimensional diffusions; a careful analysis of the scale and speed changes involved in the construction of these RED processes, together with a skew-product representation for a pair of independent Bessel processes led the authors of [9] to the recovery of some of Professor Tanaka's results via the asymptotics of local time functionals. A synthetic discussion of Sinaï's walk via stochastic calculus is presented by Z. Shi ([10]).

To conclude, let me convey the admiration of a number of collegues in France for the remarkable homogeneity of the works of Professor Tanaka which led him from the asymptotics of one-dimensional diffusions to those of the RED processes.

References

[1] Ray, D. : Sojourn times of a diffusion process. Illinois J. Math. **7** (1963), 615-630.

[2] Knight, F. : Random walks and a sojourn density process of Brownian motion. Trans. Amer. Math. Soc. **107** (1963), 56-86.

[3] McGill, P. : Markov properties of diffusion local time : a martingale approach. Adv. Appl. Prob. **14** (1982), 789-810.

[4] Jeulin, T. : Ray-Knight's theorems on Brownian local times and Tanaka formula. Seminar on Stoch. Proc. 1983, Birkhäuser (1984), 131-142.

[5] Rogers, L.C.G. and J. Walsh : The intrinsic local time sheet of Brownian motion. Prob. Th. Rel. Fileds. **88** (1991), 363-379.

[6] Rogers, L.C.G. and J. Walsh : Local time and stochastic area integrals. Ann. Prob. **19** (1991), 457-482.

[7] Yamada, T. : Principal values of Brownian local times and their related topics. In : N. Ikeda, S. Watanabe, M. Fukushima, H. Kunita (eds). Itô's stochastic calculus and probability theory. Springer (1996), 413-422.

[8] Rosen, J. and M. Yor : Tanaka formulae and renormalization for triple intersections of Brownian motion in the plane. Ann. Prob. **19** (1991), 142-159.

[9] Hu, Y., Z. Shi and M. Yor : Rates of convergence of diffusions in drifted Brownian potentials. Trans. Amer. Math. Soc. **351** (1999), 3915-3934.

[10] Shi, Z. : Sinaï's random walk via stochastic calculus. To appear in "Milieux Aléatoires", Collection : Panorama et Synthèse (Soc. Math. France).

Contributions and Influences of Professor Tanaka in Stochastic Analysis

by Shinzo Watanabe

The aim of this note is to express my hearty admiration and respect to Professor Hiroshi Tanaka for his works in probability theory which have influenced much and are playing important roles in its modern developments. As we can see by his great contributions collected in this volume, his works have been widely ranged in the field of stochastic analysis with two high peaks around main themes; stochastic models associated with nonlinear equations, particularly Boltzmann equations, and diffusion models in random environment. For me it is difficult and also, I think I am not well qualified, to make comments on his works in these main themes. I would rather give some historical account, combined with my own experience, of his contributions and influences in the area of stochastic analysis and the theory of stochastic differential equations (SDE's) in general. I should say, however, that motivations of Professor Tanaka must have been always to challenge new and unsolved problems; developing or organizing general theories may not have been his major concern.

Around 1960, when I started studying the theory of stochastic processes as a graduate student at Kyoto University, Professor Tanaka was a Research Associate at Kyushu University and was just making his debut as a young researcher in this area. At that time, people were interested in and studied extensively the theory of Markov processes, as was initiated and developed by Kolmogorov and Feller, among others. For example, my teacher, Professor K. Itô had been working jointly with H. P. McKean on sample paths description and construction of one-dimensional diffusion processes which attracted much attention of a group of young probabilists in Japan. Certainly, Itô-McKean's theory is a really important and beautiful achievement in probability theory. On the other hand, we know now well that the essential part of Itô's theory of SDE's had been established by that time and his AMS Memoir on the SDE theory was already published in 1951. However, students of Itô at that time could have had few occasions of learning the SDE theory directly from him at Kyoto University. Rather, I came to know Itô's theory through the guidance of several members in the group of Japanese probabilists like N. Ikeda, M. Nisio and Tanaka.

Tanaka's formula for the local time of one-dimensional Brownian motion was my first exciting experience of usefulness of Ito's stochastic calculus. As a graduate student, I read, in a seminar of Professor Itô, Trotter's paper [T] which established for the first time the existence of Brownian local time. Also, he kindly showed me a part of the manuscript which he was jointly preparing with McKean (this now appeared as famous Itô-McKean's book [IM]) in which was given a proof of Trotter's theorem. Both Trotter's paper and Itô-McKean's manuscript were very difficult for me. However, Brownian local time plays a key role in the Itô-McKean construction of sample functions for one-dimensional diffusion processes so that it was necessary to have some deep understanding of Trotter's theorem. A few years later, when I was still a graduate student, Professor Itô was abroad and Professor Ikeda was our adviser and sometimes our joint collaborator, Ikeda got a letter from Tanaka who was then visiting McKean at MIT, and Ikeda informed us that Tanaka found a nice proof of Trotter's theorem by applying the notion of Itô's stochastic integrals. Of course, we could not completely understand what exactly it was judging from the little information. On the other hand, Ikeda and I knew and read a paper by Skorohod [S] for reflecting Brownian motion on a one-dimensional interval in which local times at boundary points play an important role so that we could immediately have a vague idea what it might be. Later, McKean's book [M] appeared and I could then finally and completely understand the splendid idea of Tanaka. Skorohod's paper regarded Tanaka's formula as a stochastic equation and established its uniqueness of solutions. This is the origin of *Skorohod equations* to which, we know, Tanaka has also made an essential contribution ([24] in this volume).

Now Tanaka's formula is an indispensable content in every standard text book of stochastic analysis. It is now defined most generally for any continuous semimartingale. I have to confess that, at first, I could not well see the importance of such a generalization. Gradually, however, through the work of Le Gall [L] who applied generalized Tanaka's formula to the pathwise uniqueness problem for solutions of SDE's and also through more recent works in [BEKSY] and [EY] which discuss Tsirelson's discovery that the natural filtration of Walsh's Brownian motion on $n \geq 3$ rays cannot be a Brownian filtration, for example, I came to recognize that such a generalization (as treated nicely in the text books by Revuz-Yor [RY] or Rogers-Williams [RW], among others) is really important and useful.

Tanaka's formula may be considered an extension of Itô's formula so that it may be reasonably called Itô-Tanaka's formula, as well. An important and interesting aspect of this formula is that it *produces a new process*, namely, the local time: If we consider a similar extension of Itô's formula, we may obtain another interesting new process. An important idea to extend Itô's formula along the line of Tanaka's formula was given by M. Fukushima in his theory of Dirichlet forms and symmetric Markov processes associated with them. He introduced a class of stochastic processes *with zero energy* and extended Itô's formula by using this notion. Important examples of new processes thereby obtained contain the Cauchy principal value of the Brownian local times introduced and studied by T. Yamada and M. Yor, among others (cf. [Ya], [Yo]).

As I remarked above, Itô's theory on SDE's has already been established around 1960 but the only books available to study this theory at that time were those by Doob [D] and Itô [I 1], [I 2]. These are most valuable books which still keep high values at present. Fortunately in Japan, there was still another valuable research monograph written by Tanaka jointly with M. Hasegawa ([TH]) in 1964. It was one volume (Vol.19) in the series of research monographs written in Japanese, printed mimeographically and published by Seminar on Probability, an organization formed voluntarily by younger members of probabilists in Japan. It was a time when the communications and circulations were much less convenient than the present and these research monographs, some of which contained top class results at the world research level, have not been known to the world outside Japan. These promoted so much the progress of the probability theory in Japan, particularly, they helped those who just began to study in this field. For me, it was really lucky that I came across this monograph by Tanaka right after I had learned, from books by Doob and Itô, the basic elements of Itô's theory such as stochastic integrals, Itô's formula, SDE's with Lipschitz continuous coefficients and diffusion processes associated with them, and so on. It taught me, among others, the importance of *studying separately* the uniqueness problem and the existence problem for solutions of SDE's. In this monograph, Tanaka discussed SDE's with continuous coefficients; the existence of solutions was established by following Skorohod's method; for the uniqueness problem, he only wrote that it is an interesting open problem whether the solutions are unique or whether there exist solutions which define a Markov process. However, he could immediately solve ([8] in this volume) the question concerning the existence of Markovian solutions in the nondegenerate case. This Tanaka's result was further improved by D. W. Stroock and S. R. S. Varadhan who established the law uniqueness of solutions in the nondegenerate case, and by N. V. Krylov who obtained a Markov selection theorem for solutions in the general case (cf. [SV] for these results).

As far as the theory of Markov processes is concerned, what is necessary is only the existence of weak solutions and their uniqueness in law which is also equivalent to the existence of solutions and the well-posedness of martingale problems. However, in the study of stochastic flows generated by SDE's and in the innovation problems (the problems concerning filtrations and noises) associated with solutions, for example, the notion of strong solutions and relevant notion of pathwise uniqueness are also important. The terms like *strong solutions, weak solutions* and *pathwise uniqueness* did not appear explicitly in this monograph but I think that Tanaka might have been the first in Japan who recognized the importance of these notions. Actually, Tanaka is the first to give a simple and most understandable example (known as *Tanaka's example*, cf. e.g., [IW], [RW]) of SDE which enjoys uniqueness in law but *no* pathwise unique solutions. In this monograph, the pathwise uniqueness of one-dimensional SDE's with Hölder continuous coefficients of order $\alpha > 1/2$ was obtained. Later, when I was working on continuous state branching processes, I needed some uniqueness result for SDE's with Hölder continuous coefficients of order $\alpha = 1/2$ and I suddenly remembered the proof of Tanaka in this monograph. By modifying his proof, I could manage to obtain the necessary result. The proof of pathwise uniqueness I gave was still not general enough but Yamada found a nice improvement of Tanaka's proof so that an almost best possible condition for the pathwise uniqueness was obtained. However, it was not obvious for us that the pathwise uniqueness implies the law uniqueness so that it can define a strong Markov process, and we could not find any existing literature concerning this problem. Such were our main motivations of joint work [YW] and we owe much to Tanaka in this joint work.

Among many important contributions and influences of Tanaka in connection with SDE's, I would choose one more topic. This is concerned with his SDE based on a Poisson point process, which he introduced to describe a stochastic model associated with Boltzmann's equation of Maxwellian gas without cutoff ([16] in this volume). The notion of Poisson random measures (Poisson point processes) has been established by Itô, among others, in the study of the Lévy-Itô theorem for the paths structure of additive processes and a general theory of SDE's based on them has been discussed by Itô, Gihman and Skorohod, among others. It seems to me, however, that there had not been any significant application of the theory before Tanaka: I believe that Tanaka's equation is the first truly significant example of SDE's based on Poisson point processes. I remember well when Tanaka visited us and gave a series of lectures on this topic at Kyoto University around 1970. I was really fascinated by his splendid way of defining and applying a SDE based on a cleverly chosen Poisson point process. Later, I discussed with S. Takanobu on some application of SDE's based on Poisson point processes of Brownian excursions to solve SDE's for diffusion processes in a domain satisfying the most general Wentzell's boundary conditions. The idea was motivated by the work of Tanaka and I remember I was always wishing to imitate Tanaka's method, even in a small way, in this problem.

References

[BEKSY] M. T. Barlow, M. Émery, F. B. Knight, S. Song et M. Yor, Autour d'un théorème de Tsirelson sur des filtrations browniennes et non browniennes, *Sém. Prob. XXXII*, Lecture Notes in Math., **1686**, Springer, Berlin (1998), 264-305

[D] J. L. Doob, *Stochastic Processes*, John Wiley & Sons, New York, 1953

[EY] M. Émery et M. Yor, Sur un théorème de Tsirelson relatif à des mouvements browniens corrélés et à la nullité de certains temps locaux, *Sém. Prob. XXXII*, Lecture Notes in Math., **1686**, Springer, Berlin (1998), 306-312

[IW] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, Second Edition, North-Holland/Kodansha, Amsterdam/Tokyo, 1988

[I 1] K. Itô, *Kakuritsuron* (*Probability Theory*), in Japanese, Iwanami-Shoten, Tokyo, 1955

[I 2] K. Itô, *Lectures on Stochastic Processes*, Tata Institute of Fundamental Research, Bombay, 1961

[IM] K. Itô and H. P. McKean, *Diffusion Processes and their Sample Paths*, Springer, Berlin, 1965, 1974-Printing in *Classics in Mathematics*, 1996

[L] J. F. Le Gall, Applications du temps local aux équations différentielles stochastiques unidimensionnelles, *Sém. Prob. XVII*, Lecture Notes in Math., **986**, Springer, Berlin (1983), 15-31

[M] H. P. McKean, *Stochastic Integrals*, Academic Press, New York,1969

[RY] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion,* Third Edition, Springer, Berlin, 1999

[RW] L.C.G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales, Vol. 2: Itô's Calculus*, John Wiley & Sons, New York, 1987

[S] A. V. Skorohod, Stochastic equations for diffusion processes in a bounded region, Theor. Prob. Appl., **6** (1961), 264-274

[SV] D. W. Stroock and S. R. S. Varadhan, *Multidimensional Diffusion Processes*, Springer, Berlin, 1979

[TH] H. Tanaka and M. Hasegawa, *Kakuritsu Bibun Hōteishiki,* (*Stochastic Differential Equations*), in Japanese, Seminar on Prob. **17**, 1964

[T] H. F. Trotter, A properties of Brownian motion paths, Illinois J. Math. **2** (1958), 425-433

[Ya] T. Yamada, Principal values of Brownian local times and their related topics, in *Itô's Stochastic Calculus and Probability Theory, ed. N. Ikeda et al.*, Springer, Tokyo (1996), 413-422

[YW] T. Yamada and S. Watanabe, On the uniqueness of solutions of stochastic differential equations, Jour. Math. Kyoto Univ. **11** (1971), 155-167

[Yo] M. Yor, *Some Aspects of Brownian Motion,* Part II, Some Recent Martingale Problems, Lectures in Math. ETH Zürich, Birkhäuser, Basel, 1997

Some Comments on My Mathematical Works in Retrospect

by Hiroshi Tanaka

In this short note I would like to give some comments on my works on probability theory, first by mentioning briefly the initial stage of my mathematical development and then by concentrating on those works which are related mainly to Boltzmann's equation and diffusion processes in random environments.

When I was a graduate student at Kyushu University (1955–1958), the theory of Markov processes with related potential theory went through much development. As a student of Professor Gisiro Maruyama I began to study probability theory toward the above field by reading papers on Markov processes by A.N. Kolmogorov, W. Feller, J.L. Doob, K. Itô and E.B. Dynkin. I also learned many things from N. Ikeda, T. Watanabe and T. Ueno, who also began to work in this field and were taking the initiative in research activities of young probabilists in Japan, in particular, in the Probability and Statistics Summer Seminar, the April Probability Symposium and the publication of Seminar on Probability (in Japanese). I was encouraged by Ikeda through frequent conversation; I owed much to Watanabe for the problem of the paper [1] † ; the motivation of the paper [4] was to realize Ueno's idea on invariant measures of recurrent Markov processes ([U1]).

In 1957–1958 H.P. McKean visited Kyoto University, where K. Itô was at the time. He gave interesting lectures there, sometimes in other places as well, and inspired young probabilists in Japan. Itô and McKean were developing their theory on one-dimensional diffusion processes which is now found in the book [IM]. I learned some of the essential points directly by listening to their lectures and also through Itô's seminar which I sometimes attended. By analogy with their results in one-dimension Itô and McKean introduced a class of multi-dimensional diffusion processes $X(t)$ admitting the same system of harmonic measures as Brownian motion $B(t)$ and anticipated the construction of such $X(t)$ by means of time-change from $B(t)$. I carried out the actual construction but under some condition on mean exit times of $X(t)$ from bounded domains (I arrived at additive functionals of Brownian motion as a multi-dimensional substitute, though less explicit, of one-dimensional local time integrals due to Itô-McKean). McKean also obtained almost the same result. The final result was published as the

†Numbers in brackets refer to the Bibliography of Hiroshi Tanaka (see pages 425–428)

joint paper [5] which was written up while I was visiting MIT (1960–1962) at the invitation of McKean. I was much influenced by him and afterward became interested in Boltzmann's equation through his paper [M].

After writing the paper [8] on the existence of diffusion processes with continuous coefficients, I was interested in the approach from stochastic differential equations (SDEs). The uniqueness result which was lacking in my paper [8] was given later by D.W. Stroock and S.R.S. Varadhan [SV] in a clear way by introducing martingale problems. I had some misunderstanding on (or expected too much from) the pathwise uniqueness in SDEs and thus obtained no results, but from the works of A.V. Skorohod (Chapter 3 of the book [Sk]) and I.V. Girsanov ([G]) I was able to acquire some knowledge on strong and weak solutions of SDEs and often used SDEs in my later works.

In the research of probability theory in Japan around the mid 1960s there was a trend of taking up some nonlinear problems from the view-point of probability theory; in particular, N. Ikeda, M. Nagasawa and S. Watanabe were discussing nonlinear problems in connection with branching processes in a series of interesting papers (e.g. [INW]). I myself attempted to treat quasilinear parabolic equations making use of SDEs ([10]) but it had no connection with concrete problems. In 1966 McKean ([M]) introduced a class of Markov processes (with nonconstant transition mechanism) that can be associated with certain nonlinear parabolic equations such as Burgers' equation and Boltzmann's equation. I soon became interested in [M]. Probably the willingness to study Boltzmann's equation had also come from my experience of reading R. Kubo's book "Statistical Mechanics" (published in 1952, Japanese) when I was a sophomore at Kyushu University. By reading [M] carefully and comparing it with Kac's earlier paper [K], I could understand its main part: The spatially homogeneous Boltzmann's equation (nonlinear) can be derived from the master equation (linear) through the propagation of chaos and this idea is applicable to certain nonlinear parabolic equations. But the rigorous verification had been carried out only in few cases such as Boltzmann's equation of Maxwellian and cutoff type by using an analytic method. I wanted to discuss the problem by a more probabilistic method. It was easy to write down the SDE (based on a Poisson random measure) describing the Markov process associated with spatially homogeneous Boltzmann's equation but the actual proof of the existence and uniqueness of solutions was usually hard and interesting particularly when the total scattering cross section (t.s.c.s. for short) is infinite. The papers [16] and [22] treated the case of Maxwellian and non-cutoff type. The idea behind was in an extension of the previous work[15] on Kac's one-dimensional model of Maxwellian molecules in which a sort of mean square deviation from the Gaussian distribution was effectively used. In solving the SDE there was a difficulty due to the infiniteness of t.s.c.s. but the Maxwellian type provided a special method, so to say, an iteration method in the sense of law in solving the SDE (this was not applicable to Boltzmann's equation in general).

Since the argument of [22] seemed technically complicated, I tried to simplify it with some improvement on the SDE in [38]. Let (Ω_1, P_1) be a *copy* of the basic probability space (Ω, P) on which a solution $X(t,\omega)$ is to be constructed. Then the SDE (written in the form of an integral equation) was improved to

$$X(t,\omega) = X(0,\omega) + \int_{(0,t]\times\Lambda\times\Omega_1} b(X(s-,\omega), X(s-,\omega_1), \lambda) N(dsd\lambda d\omega_1),$$

where $N(\cdot)$ is a Poisson random measure on $(0,\infty)\times\Lambda\times\Omega_1$ with intensity measure $ds\mu(d\lambda)P_1(d\omega_1)$, (Λ, μ) being a suitable measure space. The presence of the copy $X(s-,\omega_1)$ of an unknown $X(s-,\omega)$ matches the *nonlinearity* of the problem.

I thought that there was a gap between the uniqueness of solutions to the SDE of [16] and the uniqueness of solutions to the Boltzmann equation itself, so I wrote the paper [19] to fill this gap but later was able to obtain a better result, namely the uniqueness of probability measure solutions of the Boltzmann equation itself (unpublished) by making use of Varadhan's result ([V]) concerning the uniqueness of probability measure solutions of certain parabolic equations.

When writing the paper [6] with K. Sato, in 1962, which was motivated by Ueno's idea on Markov processes on boundaries ([U2]), I became aware that a reflecting Brownian motion $X(t)$ for a multi-dimensional domain can be expressed as $B(t) + \Phi(t)$ where $B(t)$ is a standard Brownian motion and $\Phi(t)$ is an integral with respect to the local time on the boundary for $X(t)$. However, I could not regard this expression as an equation in which $X(t)$ and $\Phi(t)$ are considered as unknown, because I did not anticipate the uniqueness of solutions at that time. Later I noticed that the uniqueness can be obtained easily from the convexity assumption on the domain, and so I could formulate Skorohod's problem for convex domains ([24]). The convexity assumption was soon removed by P.L. Lions and A.S. Sznitman ([LS]). Thanks to their result I was able to construct the motion of a finite number of mutually reflecting Brownian balls as a solution to Skorohod's problem in [34] (the infinite case was discussed later by H. Tanemura [T]).

In 1984 I stayed in Zürich for three months at the invitation of H. Föllmer and Nagasawa. I attended the probability seminar organized jointly by Föllmer (ETH Zürich) and Nagasawa (the University of Zürich), which was fresh and active, with many young probabilists. This visit gave much influence to my later works. A series of joint works with Nagasawa ([31][33][39][42]) began around this time; there we discussed some types of many particle systems with mean field interaction which arose from Nagasawa's idea on the Schrödinger equation.

At that time T. Brox, a postdoctral fellow at the University of Zürich, was working on a diffusion analogue (appeared in [B]) of Ya.G. Sinai's random walk in a random environment ([S]). I came to know the results of Sinai and Brox for the first time (as I knew later there is also S. Schumacher's work [Sch] on a similar subject) and was much impressed by their works since it seemed to me that the results gave a new type of limit theorems involving a scaling quite different from those I had ever encountered. I thus became interested in stochastic processes in random environments.

A diffusion process in a Brownian environment or Brox's diffusion for simplicity is a continuous time version of Sinai's random walk in a random environment and described by the (symbolic) SDE

$$dX(t) = dB(t) - \frac{1}{2}w'(X(t))dt, \qquad X(0) = 0,$$

where $B(t), t \geq 0$, is a one-dimensional Brownian motion independent of a Brownian environment $w(x)$, $x \in \mathbb{R}$. For a precise definition, e.g. see [B] or [41]. Brox's diffusion is sometimes called a Brownian motion in a white noise environment. The arguments of Brox were based on many of Sinai's ideas ([S]), among which the notion of a valley — a depression in English translation ([S]) — played an essential role together with the self-similarity of the random environment. Reading Brox's preprints I was able to obtain a better understanding of Sinai's result. Since I preferred to work with Brownian motion rather than with random walks, I considered most of the problems in the framework of diffusion processes. Most of the results obtained for diffusions in random environments were preceded by the corresponding ones for random walks; however, working with diffusion processes was more fun for me; it often seemed to clarify methods and results since one could use a lot of fine properties of Brownian motion.

I give some comments on my works concerning the following problems.

(A) Some refinements of the results of [S][B][Sch] (localization).

(B) A generalization of the problem (A) to a wider class of random environments.

(C) A Brownian motion with drift in a white noise environment.

(D) A multi-dimensional problem.

(A): When I heard from Brox that $(\log t)^{-2}X(t) - b(t)$ tends to 0 in probability as $t \to \infty$ (where $b(t)$, for each fixed t, is defined as the bottom of a suitable valley and admits a probability distribution independent of t), I got interested in the limit theorem of the next stage: Can one obtain a non-degenerate limit distribution by a suitable rescaling of $(\log t)^{-2}X(t) - b(t)$? I could not immediately anticipate that the multiplication by $(\log t)^2$ simply is enough to have a limit distribution. The paper [41] discusses this problem. The result is that $X(t) - (\log t)^2 b(t)$ admits a non-degenerate limit distribution as $t \to \infty$ which can be expressed in terms of independent Bessel processes β_+ and β_- appearing in (ii) below. This result seemed interesting particularly when compared with the fact that an individual diffusion process with frozen environment w is null-recurrent (a.s.). I must remark that there is also an earlier work [Go] by A.O. Golosov on the same subject for random walks. Following [Go] and also for terminological convenience I call a result of this type *localization* (in some occasions the term *localization* is used in a wider sense). The key points of the argument in [41] were (i) to refine Brox's method and (ii) to describe the local behavior (law) of the environment around the bottom $b(e)$ in terms of two independent Bessel processes β_+ and β_- of index 3. This part, (ii), was most interesting for me, in which I used the construction of a reflecting Brownian motion on $[0, \infty)$ by means of the associated Poisson point process of excursions ([I][IW:pp.123–125]) and then a certain reversal of time. Theorems of D. Williams [W] and J.W. Pitman [P] were also used. The same problem as [41] was again discussed in [52], in which the part (i) was replaced by a use of a theorem of Y. Ogura ([O], see also p.756 of [52]) and consequently the whole story of the proof became transparent.

(B): In [47] we proved that the localization of the type discussed in [41] holds for a certain class of random environments that are asymptotically strictly stable. But after having written [52] I began to anticipate that the localization will hold if we take for the random environment $\{w(x), x \in \mathbb{R}\}$ any Lévy process satisfying the following condition: There exists $\alpha\,(0 < \alpha < 2)$ such that every positive sequence $\{\lambda_k\}$ tending to ∞ contains a subsequence $\{\lambda'_k\}$ along which $\{\lambda^{-1}w(\lambda^\alpha x), x \in \mathbb{R}\}$ converges in law to some nice non-degenerate process *that may depend on* $\{\lambda'_k\}$. So I wanted to find a *good* class of Lévy random environments for which the localization of the type discussed in [41] holds. For this purpose I was often led to the study of some properties of Lévy processes themselves and this was also fun for me. For instance, I studied superharmonic transforms of absorbing Lévy processes in the half line $(0, \infty)$ in order to grasp the asymptotic law of a certain part of the environment around some bottom (a summary in Japanese is in Cooperative Research Report 51, The Institute of Statist. Math.; what I had for the proof of (ii) of Theorem 7 therein is incomplete).

(C): The limiting behavior of the process $X(t)$ described by

$$dX(t) = dB(t) + \frac{\kappa}{2}dt - \frac{1}{2}w'(X(t))dt$$

can not be treated by the method of Sinai-Brox-Schumacher. In 1986 when I was at Minnesota, B. Tóth pointed out to me that this would correspond to the random walk model discussed in an earlier work [KKS] by H. Kesten M.V. Kozlov and F. Spitzer. However, an adaptation of their method to our case seemed very hard to me. On the other hand, it was already recognized that Krein's spectral theory of strings (e.g. see [KK]) provides a powerful method in the study of various problems on one-dimensional diffusion processes (in non-random environments); for example, see S. Kotani-S. Watanabe [KW] and the references therein. In 1989 K. Kawazu and I had a chance, at a small symposium, to discuss with Kotani on the problem of limiting behavior of the above $X(t)$ from the view point of the applicability of Krein's theory and learned the method of Laplace transform of a hitting time, i.e., Kotani's formula (unpublished, see [56]), which he had obtained previously in finding the limit distribution of the maximum of the Brox-Schumacher diffusion. We thus became interested in using Kotani's formula and Krein's theory for our problem and were able to obtain the first result [56].

But it gave only a partial answer; namely, the case $1 \le \kappa \le 2$ was not discussed, in which the limit distribution was expected to have the whole real line for the support. So I was misled, for a good while, to a false conclusion that the method of Laplace transform (i.e., Kotani's formula) would not work. However, luckily it worked as was discussed later in the second paper [57] where, in addition to [KK], I also used Krein's more specific result [Kr] (see also [DM:p.266] [KW:p.239]). The same result was obtained also by Y. Hu, Z. Shi and M. Yor ([HSY]) by a method quite different from ours. Finally I remark that an environment-wise central limit theorem was proved in the case $\kappa > 2$ (see [55]); the proof needed a series of detailed computations in spite of the easy-looking problem. I do not know whether a similar result holds for $\kappa = 2$. (The right hand side of the equality in the 13th line of p.1809 of [57] should be

$$(2^7\pi)^{-1/2} \int_{-\infty}^{t} \exp\{-2^{-7}s^2\}ds.$$

I remark also that the formula at the top of p.201 of [36] is not correct; it was corrected in [35].)

(D): P. Mathieu ([Ma]) considered a diffusion process in a multi-dimensional Brownian environment. The paper [51] proved its recurrent property. The process $X(t)$ corresponds to one of random walk models in R. Durrett [D], but is not a diffusion analogue of the random walk of J. Bricmont and A. Kupiainen [BK], a multi-dimensional random walk corresponding to Sinai's one. From [Ma] and [D] we see that the process $X(t)$ has the localization property of the type discussed in [S] but the existence problem of a limit distribution of $(\log t)^{-2}X(t)$ seemed unsettled. In this connection I studied some properties of sample functions of a Lévy's Brownian motion with a multi-dimensional time (e.g. [61]).

References

[BK] J. Bricmont and A. Kupiainen, Random walks in asymmetric random environments, Comm. Math. Phys. **142** (1991), 345-420.

[B] T. Brox, A one-dimensional diffusion process in Wiener medium, Ann. Probab. **14** (1986), 1206-1218.

[D] R. Durrett, Multidimensional random walks in random environments with subclassical limiting behavior, Comm. Math. Phys. **104** (1986), 87-102.

[DM] H. Dym and H.P. McKean, Gaussian Processes, Function Theory, and the Inverse Spectral Problem, Academic Press, 1976.

[G] I.V. Girsanov, An example of nonuniqueness for solutions of K. Itô's stochastic integral equations, Theor. Probab. Appl. **7** (1962), 336-342.

[Go] A.O. Golosov, Localization of random walks in one-dimensional random environments, Comm. Math. Phys. **92** (1984), 491-506.

[HSY] Y. Hu, Z. Shi and M. Yor, Rate of Convergence of diffusions with drifted Brownian potentials, Trans. Amer. Math. Soc. **351** (1999), 3915-3934.

[I] K. Itô, Poisson point processes attached to Markov processes, Proc. 6th Berkeley Symp. Math. Statist. Probab. **III**, 1970, 225–239.

[IM] K. Itô and H.P. McKean, Diffusion Processes and Their Sample Paths, Springer, 1965.

[INW] N. Ikeda, M. Nagasawa and S. Watanabe, Branching Markov processes I, II, III, J. Math. Kyoto Univ. **8** (1968), 233–278, 365–410; **9** (1969), 95–160.

[IW] N. Ikeda and S. Watanabe, Stochastic Differential Equations and Diffusion Processes, North-Holland, 1981.

[K] M. Kac, Foundations of kinetic theory, Proc. 3rd Berkeley Symp. Math. Statist. Probab., 3, 1955, 171–197.

[Kr] M.G. Krein, On some cases of the effective determination of the density of a non-homogeneous string from its spectral function, Dokl. Akad. Nauk SSSR **93** (1953), 617–620.

[KK] I.S. Kac and M.G. Krein, On the spectral functions of the string, Amer. Math. Soc. Transl. (2), **103** (1974), 19–102.

[KKS] H. Kesten, V. Kozlov and F. Spitzer, A limit law for random walk in a random environment, Compositio Math. **30** (1975), 145-168.

[KW] S. Kotani and S. Watanabe, Krein's spectral theory of strings and generalized diffusion processes, 235-259 in Lecture Notes in Math. **923**, Springer, 1982.

[LS] P.L. Lions and A.S. Sznitman, Stochastic differential equations with reflecting boundary conditions, Comm. Pure Appl. Math. **37** (1984), 511–537.

[Ma] P. Mathieu, Zero white noise limit through Dirichlet forms, Application to diffusions in a random medium, Probab. Theory Related Fields **99** (1994), pp. 549-580.

[M] H.P. McKean, A class of Markov processes associated with nonlinear parabolic equations, Proc. Nat. Acad. Sci. U.S.A. **56** (1966), 1907-1911.

[O] Y. Ogura, One-dimensional bigeneralized diffusion processes, J. Math. Soc. Japan **41** (1989), 213-242.

[P] J.W. Pitman, One-dimensional Brownian motion and three-dimensional Bessel process, Adv. Appl. Probab. **7** (1975), 511-526.

[Sch] S. Schumacher, Diffusions with random coefficients, pp.351-356 in: Particle Systems, Random Media and Large Deviations, R. Durrett, ed., Contemporary Math. **41**, 1985.

[S] Y.G. Sinai, The limiting behavior of a one-dimensional random walk in a random medium, Theory Probab. Appl. **27** (1982), 256-268.

[Sk] A.V. Skorokhod, Studies in the Theory of Random Processes, Addison-Wesley, 1965.

[SV] D.W. Stroock and S.R.S. Varadhan, Diffusion processes with continuous coefficients, I, II, Comm. Pure Appl. Math. **22** (1969), 345–400, 479–530.

[T] H. Tanemura, A system of infinitely many mutually reflecting Brownian balls in $\mathbb{R}^d$, Probab. Th. Rel. Fields **104** (1996), 399–426.

[U1] T. Ueno, On recurrent Markov processes, Kōdai Math. Sem. Rep. **12** (1960), 109-142.

[U2] T. Ueno, The diffusion satisfying Wentzell's boundary condition and the Markov process on the boundary I, II, Proc. Japan Acad. **36** (1960), 533–538, 625–629.

[V] S.R.S. Varadhan, Invariant measures of Markov processes, Lecture at Kyoto Univ., 1980; A brief summary of it taken and written up by S. Kotani in Japanese is in Sugaku Vol.**33**, No.3, 1981, p.276.

[Wi] D. Williams, Path decomposition and continuity of local time for one-dimensional diffusions, I, Proc. London Math. Soc. (3) **28** (1974), 738-768.

MEMOIRS OF THE COLLEGE OF SCIENCE, UNIVERSITY OF KYOTO, SERIES A
Vol. XXXIII, Mathematics No. 3, 1961.

Additive functionals of the Brownian path

By

H. P. MCKEAN, JR. and Hiroshi TANAKA

(Communicated by Prof. K. Itô, March 14, 1961)

1. INTRODUCTION.

A functional $\mathfrak{f}(t)=\mathfrak{f}(t, w)$ of $t(\geq 0)$ and a several-dimensional Brownian path $w: t\to x(t)$ is said to be *additive* if

1.1 $\mathfrak{f}(t, w)$ *depends upon* t *and* $x(s): s\leq t$ *alone.*
1.2 $0=\mathfrak{f}(0)\leq \mathfrak{f} < +\infty$
1.3 $\mathfrak{f}(t\pm)=\mathfrak{f}(t)$
1.4 $\mathfrak{f}(t)=\mathfrak{f}(s)+\mathfrak{f}(t-s, w_s^+) \qquad t\geq s$,

where w_s^+ is the shifted path $w_s^+: t\to x(t+s)$; for example, $\mathfrak{f}(t)=\int_0^t f(x(s))ds$ is an additive functional if $f\geq 0$ is bounded and Borel.

K. Itô and H. P. McKean, Jr. [13] proved that in the 1-dimensional case such an additive functional is an *integral*

1.5 $$\mathfrak{f}(t)=\int \mathfrak{t}(t, b)e(db)$$

of the *standard Brownian local times*

1.6 $$\mathfrak{t}(t, a)=\lim_{b\downarrow a}\frac{\text{measure }(s: a\leq x(s)<b, s\leq t)}{b-a}$$

with respect to a non-negative measure e, finite on bounded intervals.

Brownian local times are not available in $d\geq 2$ dimensions, but $\mathfrak{f}$ can still be expressed as a (*formal*) Hellinger *integral*

1.7 $$\mathfrak{f}(t)=\int\frac{\text{measure }(s: x(s)\in db, s\leq t)e(db)}{db}$$

with a non-negative measure e which is *smooth* in the sense that *each bounded open D is the union of an increasing series of sets* $B_n(n\geq 1)$, *closed in D, such that*

1.8a *the charge distributions* $e|B_n$ *have bounded potentials* $\int_{B_n} G de \leq n(n\geq 1)$, *where G is the Green function of D;*

1.8b *for large n, depending upon the path, the Brownian particle lies in* B_n *until it leaves D, i.e.,*

$$P_\bullet[x(t)\in B_n,\ t<\min(s: x(s)\notin D),\ n\uparrow +\infty] = 1\,,$$

where $P_\bullet(B)$ *is the probability of the event B as a function of the starting point of the Brownian motion.*

The correspondence embodied in 1.7 *between the class of smooth measures e and the class of additive functionals* $\mathfrak{f}$ *is one to one and onto.*

1.8a implies that $e(B)=0$ unless B has positive logarithmic capacity in the 2-dimensional case or has positive Newtonian capacity in the $d\geq 3$ dimensional case; thus, a smooth measure cannot attach positive mass to a line in 3 dimensions, nor, in 4 dimensions, to a surface. But it can be singular relative to Lebesgue measure; the simplest example in 3 dimensions is the uniform distribution on the spherical surface $|a|=1$.

Choose $d=3$, $e(db)=|b|^\alpha db$, and let G be the Green function of D: $|a|<1$; then

1.9a
$$\begin{aligned} p = \int G de & \\ \leq \text{constant } (2+\alpha)^{-1} & \qquad \alpha > -2 \\ < +\infty \text{ except at } 0 & \qquad -3<\alpha\leq -2 \\ \equiv +\infty & \qquad \alpha\leq -3\,, \end{aligned}$$

1.9b
$$\mathfrak{E}(e) \equiv \iint G de de < +\infty \qquad \alpha > -5/2\,,$$

and it follows that e is smooth for $\alpha>-2$. *e is not smooth for* $\alpha\leq -2$; in fact, choosing $B_n\uparrow D$ as needed for 1.8,

1.10
$$\begin{aligned} & E_0\left[\int_0^\varepsilon |x(s)|^\alpha ds,\ x(t)\in B_n,\ t<\varepsilon\right] \\ & \leq \int_0^{+\infty} ds \int_{B_n} P_0[x(s)\in db,\ \max_{t\leq s} |x(t)|<1]|b|^\alpha \\ & \leq \int_{B_n} G(0,b) de < +\infty\,, \end{aligned}$$

thanks to 1.8a, and

1.11 $\lim_{\varepsilon\downarrow 0}\lim_{n\uparrow+\infty} P_0[x(t)\in B_n,\ t<\varepsilon]=1$

thanks to 1.8b, while, as is not hard to prove,

1.12 $P_0\left[\int_0^\varepsilon |x(s)|^\alpha ds \equiv +\infty\ ,\ \varepsilon>0\right]=1\,,$

contradicting 1.10.

V. A. Volkonskii [15, 16] also studied additive functionals, establishing a special case of the above for a wider class of motions; the method used below is similar to his.

Given a 1-dimensional diffusion with the same hitting probabilities as the standard Brownian motion:

1.13 $P_\xi[\min(t:\ x(t)=a)<\min(t:\ x(t)=b)]$
$$=\frac{b-\xi}{b-a}\qquad a<\xi<b\,,$$

W. Feller [8] explained how to express the associated generator $\mathfrak{G}$ as a *differential opearator* based upon a *speed measure e*, positive on open intervals:

1.14 $$\mathfrak{G}u=\frac{u^+(da)}{e(da)}=\lim_{b\downarrow a}\frac{u^+(b)-u^+(a)}{e(a,b)}$$
$$u^+(a)=\lim_{b\downarrow a}\frac{u(b)-u(a)}{b-a}\,,$$

and K. Itô and H. P. McKean, Jr. [13] found that its sample paths could be expressed as standard Brownian sample paths run with the *stochastic clock* $\mathfrak{f}^{-1}$ which is the inverse function of the additive functional (local time integral) $\mathfrak{f}=\int \mathfrak{t}\,de$ associated with the speed measure.

V. A. Volkonskiî [15] also studied such time substitutions; his method is less explicit because it does not use local times but has the advantage that it can be applied in higher dimensions.

As will be explained below, a $d\geq 2$ dimensional *diffusion with Brownian hitting probabilities* has as its generator *the closure of a differential operator*

1.15 $$\mathfrak{G}u = -\frac{e^{u}(db)}{e(db)}$$

based upon a (smooth) *speed measure* e, positive on the neighborhoods of H. Cartan's *fine topology* [2]; moreover, the associated motion is the standard Brownian motion run with the inverse function $\mathfrak{f}^{-1}$ of the additive functional $\mathfrak{f}$ associated with e, and *this correspondence between the class of diffusions with Brownian hitting probabilities and the class of smooth measures e positive on fine neighborhoods is one to one and onto.*

2. BROWNIAN MOTION.

Choose $d \geq 2$, let $E^d = R^d$ it $d=2$, let it be the one-point compactification $R^d + \infty$ if $d \geq 3$, introduce the space of continuous *sample paths* $w : [0, +\infty) \to E^d$ with

2.1 $$\begin{aligned} w(t) &\in R^d & t &< \mathfrak{m}_\infty \\ &= \infty & t &\geq \mathfrak{m}_\infty, \end{aligned}$$

where $\mathfrak{m}_\infty = \mathfrak{m}_\infty(w) \leq +\infty$ and $\mathfrak{m}_\infty \equiv +\infty$ in case $d=2$, let $w(t) = x(t, w) = x(t)$ as need be, note that $x(+\infty) \equiv \infty$ even if $d=2$, and, introducing the corresponding *coordiante fields* $\boldsymbol{B}_t = \boldsymbol{B}[x(s) : s \leq t]$ and $\boldsymbol{B} = \boldsymbol{B}_{\infty+}$, let $P_{\cdot}(B)$ be the *probability* (*Wiener measure*) of the event $B \in \boldsymbol{B}$ as a function of the starting point of the *d-dimentional Brownian motion* with *generator* $\mathfrak{G} = \frac{\partial^2}{\partial b_1^2} + \frac{\partial^2}{\partial b_2^2} + \cdots + \frac{\partial^2}{\partial b_d^2}$.[1]

Brownian motion enthusiasts are familiar with the fact that the Brownian traveller *starts afresh* at a passage time; the full significance of this was explained by E. B. Dynkin [6] and G. Hunt [9] as follows.

An instant of time $0 \leq \mathfrak{m} \leq +\infty$ depending upon the path is said to be a *Markov time* if

2.2 $$(w : \mathfrak{m} < t) \in \boldsymbol{B}_t \qquad t \geq 0;$$

for example, the passage time $\mathfrak{m}_Q = \inf(t : x(t) \in Q)$ to a closed or

---

[1] $\mathfrak{G}/2$ is often used as the generator of the Brownian motion, but for our purpose it is simpler to omit the factor 1/2.

open d-dimensional figure Q is a Markov time and so is $\mathfrak{m}=\min\left(t: \int_0^t f(x(s))ds=1\right)$ if $0<f\leq 1$ is a Borel function.

Given such a Markov time $\mathfrak{m}$, if $w_{\mathfrak{m}}^{+}$ is the *shifted path*

2.3 $\quad w_{\mathfrak{m}}^{+}: t \to x(t+\mathfrak{m})$

and if $\boldsymbol{B}_{\mathfrak{m}+}$ is the *field* of events $B\in \boldsymbol{B}$ such that

2.4 $\quad B \cap (w: \mathfrak{m}<t) \in \boldsymbol{B}_t \qquad t\geq 0,$

then the Brownian particle *starts afresh* at time $t=\mathfrak{m}$, i.e.,

2.5 $\quad P_a[w_{\mathfrak{m}}^{+} \in B | \boldsymbol{B}_{\mathfrak{m}+}] = P_b(B) \qquad a \in E^d,\ B\in \boldsymbol{B},\ b\equiv x(\mathfrak{m})$ [2].

Blumenthal's 01 law [1]:

2.6 $\quad P_\bullet(B) = {0 \atop 1} \qquad B \in \boldsymbol{B}_{0+} = \bigcap_{\varepsilon>0} \boldsymbol{B}_\varepsilon$

is a special case of 2.5.

A. R. Galmarino[3] has pointed out that a *non-negative* Borel *function* $\mathfrak{m}$ *of the sample path is a Markov time if and only if*

2.7a $\quad \mathfrak{m}(u)<t$

2.7b $\quad x(u, u) = x(s, v) \qquad s\leq t$

imply $\mathfrak{m}(u)=\mathfrak{m}(v)$ and that *an event* $B\in \boldsymbol{B}$ *is a member of* $\boldsymbol{B}_{\mathfrak{m}+}$ *if and only if* 2.7 *coupled with* $u\in B$ *implies* $v\in B$. As a simple application of this test, note that $B_{\mathfrak{m}+}$ measures both $\mathfrak{m}$ and the *past* $x(\theta\wedge\mathfrak{m})$ $(\theta\geq 0)$[4] because 2.7 implies $\theta\wedge\mathfrak{m}(u)=\theta\wedge\mathfrak{m}(v)<t$ and hence $x(\theta\wedge\mathfrak{m}(u), u)=x(\theta\wedge\mathfrak{m}(v), v)$.

Given bounded open $D\subset R^d$ with boundary ∂D, if $\mathfrak{m}_{\partial D}$ is the *exit time* $\min(t: x(t)\in\partial D)$, then the *hitting probability*

2.8 $\quad h_{\partial D}(a, db) = P_a[x(\mathfrak{m}_{\partial D}) \in db] \qquad a\in D,\ db \in \partial D$

is the *classical harmonic measure* of db as viewed from the point a, and, if $G_D(a, b)$ is the *classical Green function* of D, then

2.9 $\quad E_a[\text{measure}\,(t: x(t)\in db, t<\mathfrak{m}_{\partial D}] = G_D(a, b)db$ [5] $\quad a, b\in D;$

---

[2] $x(\mathfrak{m})\equiv\infty$ in case $\mathfrak{m}=+\infty$; it is understood that $P_\infty[x(t)\equiv\infty,\ t\geq 0]=1$.

[3] private communication.

[4] $a\wedge b$ is the smaller of a and b.

[5] $E_\bullet(f)=\int f dP_\bullet$.

for the proofs, see J. Doob [5] and G. Hunt [9].

G. Hunt [10] has called a non-negative Borel function p *excessive* on D if

2.10 $\quad E_a[p(x(t)), t < \mathfrak{m}_{\partial D}] \uparrow p(a) \qquad t \downarrow 0, \ a \in D.$

An excessive function can be split into its *greatest harmonic minorant* h plus the *potential* $\int G_D de$ of a non-negative (Riesz) measure e, indeed, Hunt's excessive functions are the same as the *superharmonic* functions of F. Riesz [14]. J. Doob [5] proved that *an excessive function is continuous on the Brownian path* $(t < \mathfrak{m}_{\partial D})$ and that *a potential tends to* 0 *along the Brownian path* $(t \uparrow \mathfrak{m}_{\partial D})$.

3. THE ASSOCIATED MEASURE OF AN ADDITIVE FUNCTIONAL.

Consider an additive functional $\mathfrak{f}$ of the Brownian sample path as described in seition 1, interpreting 1.1 to mean

3.1 $\quad \mathfrak{f}(t, w)$ *is measurable* $\boldsymbol{B}_t$ *for each* $t \geq 0$.

The purpose of this section is *to associate with* $\mathfrak{f}$ *a unique non-negative measure* e *such that, for each bounded open* $D \subset R^d$,

3.2 $\quad 1 - p_\alpha = \alpha \int G p_\alpha de \qquad p_\alpha = E.[e^{-\alpha \mathfrak{f}(\mathfrak{m}_{\partial D})}], \ \alpha > 0,$

where G is the Green funetion of D and the integral is extended over D; *it will follow from* 3.2 *that* e *is smooth.*

Ccnsider, for this purpose, the additive functional

3.3 $\quad \mathfrak{f}_\alpha(t) = \int_0^{t \wedge \mathfrak{m}_{\partial D}} p_\alpha(x(s)) \mathfrak{f}(ds) \qquad t \geq 0,$

and let us begin with the following simple lemmas:

a) $1 - p_\alpha$ *is the potential of a non-negative measure* αe_α.

b) $E.[\mathfrak{f}_\alpha(\mathfrak{m}_{\partial D})] < +\infty$.

c) $\mathfrak{f}_\alpha \uparrow \mathfrak{f}$ *as* $\alpha \downarrow 0$.

d) $1 - p_\alpha = \alpha E.[\mathfrak{f}(\mathfrak{m}_{\partial D})] = \alpha \int G de_\alpha$.

e) $E.\left[\int_0^{\mathfrak{m}_{\partial D}} f(x(s)) \mathfrak{f}_\alpha(ds)\right] = \int G f de_\alpha$ *if* $f \geq 0$ *is a Borel function.*

f) $p_\alpha^{-1} e_\alpha(db) = e(db)$ *is independent of* α *and of* D.
g) e *is unique.*

Because

3.3 $$E_\cdot(1-p_\alpha(x(t)), t<\mathfrak{m}_{\partial D}) = E_\cdot(1-e^{-\alpha[\mathfrak{f}(\mathfrak{m}_{\partial D})-\mathfrak{f}(t)]}, t<\mathfrak{m}_{\partial D})$$
$$\uparrow 1-p_\alpha \qquad t\downarrow 0,$$

$1-p_\alpha$ is excessive; it is, in fact, a potential thanks to

3.4 $$\int_{\partial \dot{D}} h_{\partial\dot{D}}(a, db)[1-p_\alpha] = E_a(1-e^{-\alpha[\mathfrak{f}(\mathfrak{m}_{\partial D})-\mathfrak{f}(\mathfrak{m}_{\partial\dot{D}})]})\downarrow 0 \qquad \dot{D}\uparrow D,$$

and this completes the proof of a). As to b), p_α is continuous on the Brownian path because of 1, and, since $\mathfrak{f}$ is continuous, b) follows on letting $n\uparrow+\infty$ in

3.5 $$1-p_\alpha = E_\cdot\left[\int_0^{\mathfrak{m}_{\partial D}} e^{-\alpha[\mathfrak{f}(\mathfrak{m}_{\partial D})-\mathfrak{f}(t)]}\mathfrak{f}(dt)\right]$$
$$\geq \alpha E_\cdot\left[\sum_{k2^{-n}<\mathfrak{m}_{\partial D}} e^{-\alpha\mathfrak{f}(\mathfrak{m}_{\partial D}(w^+_{k2^{-n}}),\, w^+_{k2^{-n}})}e^{-\alpha\mathfrak{f}(I_k)}\,\mathfrak{f}(I_k)\right]$$
$$I_k = [(k-1)2^{-n}, k2^{-n})$$
$$= \alpha E_\cdot\left[\sum_{k2^{-n}<\mathfrak{m}_{\partial D}} p_\alpha(x(k2^{-n}))e^{-\alpha\mathfrak{f}(I_k)}\mathfrak{f}(I_k)\right].$$

Because of c), which is obvious,

3.6 $$\alpha^{-1}(1-p_\alpha) = \lim_{\varepsilon\downarrow 0} E_\cdot\left[\int_0^{\mathfrak{m}_{\partial D}} e^{-\alpha[\mathfrak{f}(\mathfrak{m}_{\partial D})-\mathfrak{f}(t)]}\mathfrak{f}_\varepsilon(dt)\right],$$

and, using the method of 3.5 and $E_\cdot[\mathfrak{f}_\varepsilon(\mathfrak{m}_{\partial D})]<+\infty$, it appears that

3.7 $$\alpha^{-1}(1-p_\alpha) = \lim_{\varepsilon\downarrow 0} E_\cdot\left[\int_0^{\mathfrak{m}_{\partial D}} p_\alpha(x(t))\mathfrak{f}_\varepsilon(dt)\right] = E_\cdot\left[\int_0^{\mathfrak{m}_{\partial D}} p_\alpha d\mathfrak{f}\right]$$

proving d).

K. Itô (private communication) pointed out the following neat method for proving e). Choose closed $B\subset D$ such that $e_\alpha(\partial B)=0$ and let e_1 and e_2 be the charge distributions of the potentials $p_1 = E_\cdot\left[\int_0^{\mathfrak{m}_{\partial D}} f d\mathfrak{f}_\alpha\right]$ and $p_2 = E_\cdot\left[\int_0^{\mathfrak{m}_{\partial D}}(1-f)d\mathfrak{f}_\alpha\right]$, in which f is the indicator function of B. Because p_1 is harmonic *outside* B and differs from $\alpha^{-1}(1-p_\alpha) = E_\cdot[\mathfrak{f}_\alpha(\mathfrak{m}_{\partial D})]$ by a harmonic function *inside* B, e_1 is not smaller than the restriction of e_α to B, and, for the same reasons, e_2 is not smaller than the restriction of e_α to $D-B$. But $p_1+p_2=\alpha^{-1}(1-p_\alpha)$, whence

$$p_1 = E.\left[\int_0^{\mathfrak{m}_{\partial D}} f d\mathfrak{f}_\alpha\right] = \int_B G de_\alpha , \tag{3.8}$$

and since such figures B generate the class of Borel subsets of D, e) follows.

As to f), $p_\alpha>0$ because $\mathfrak{f}(\mathfrak{m}_{\partial D})<+\infty$, and, choosing $0<\beta<\alpha$, e) implies

$$\int G de_\alpha = E.\left[\int_0^{\mathfrak{m}_{\partial D}} p_\alpha/p_\beta d\mathfrak{f}_\beta\right] = \int G p_\alpha/p_\beta de_\beta , \tag{3.9}$$

i.e., $de\equiv p_\alpha^{-1}de_\alpha$ is indeqendent of α; it is also independent of D because if $\dot{D}\supset D$ and if $p_\alpha\equiv 0$ outside D, then, with an obvious notation,

$$\begin{aligned}
\int G p_\alpha \dot{p}_\beta de &= E.\left[\int_0^{\mathfrak{m}_{\partial D}} p_\alpha d\dot{\mathfrak{f}}_\beta\right] \\
&= E.\left[\int_0^{\mathfrak{m}_{\partial \dot{D}}} p_\alpha d\dot{\mathfrak{f}}_\beta\right] - E.\left[\int_{\mathfrak{m}_{\partial D}}^{\mathfrak{m}_{\partial \dot{D}}} p_\alpha d\dot{\mathfrak{f}}_\beta\right] \\
&= \int \dot{G} p_\alpha \dot{p}_\beta d\dot{e} - \int dh_{\partial D} \int \dot{G} p_\alpha \dot{p}_\beta d\dot{e} \\
&= \int G \dot{p}_\alpha p_\beta d\dot{e} .
\end{aligned} \tag{3.10}$$

g) is immediate from 3.2.

To establish the *smoothness* of e, take bounded open D and put $B_n=D\cap(p_1\geq n^{-1})$. Because $1-p_1$ is a potential, B_n is closed in D and increases to D as $n\uparrow+\infty$; moreover, according to 3.2,

$$\int_{B_n} G de \leq n \int G p_1 de = n(1-p_1) \leq n , \tag{3.11}$$

which is 1.8a, and, since, along the Brownian path, $0<p_1$ is continuous and tends to 1 $(t\uparrow \mathfrak{m}_{\partial D})$,

$$\begin{aligned}
&P.[\inf(t: x(t)\notin B_n)<\mathfrak{m}_{\partial D}] \\
&= P.[\inf_{t<\mathfrak{m}_{\partial D}} p_1(x(t))<n^{-1}]\downarrow 0 \qquad n\uparrow+\infty ,
\end{aligned} \tag{3.12}$$

which verifies 1.8b.

4. UNIQUENESS.

The following simple lemma is useful in later sections: *two additive functionals* $\mathfrak{f}_1$ and $\mathfrak{f}_2$ *with the same bounded mean*

4.1 $$p = E.[\mathfrak{f}(\mathfrak{m}_{\partial D})] = \lim_{\varepsilon \downarrow 0} \varepsilon^{-1}(1-p_\varepsilon) = \int Gde \qquad \mathfrak{f} = \mathfrak{f}_1, \mathfrak{f}_2$$

are the same for $t \leq \mathfrak{m}_{\partial D}$.

An argument similar to 3.5 implies

4.2 $$\begin{aligned} &E.\left(\int_0^{\mathfrak{m}_{\partial D}} [\mathfrak{f}_i(\mathfrak{m}_{\partial D}) - \mathfrak{f}_i(t)]\,\mathfrak{f}_k(dt)\right) \\ &= E.\left(\int_0^{\mathfrak{m}_{\partial D}} p\,d\mathfrak{f}_k\right) \\ &= \int Gpde \leq \|p\|_\infty^2 < +\infty\,; \end{aligned}$$

thus, putting $\mathfrak{f} = \mathfrak{f}_2 - \mathfrak{f}_1$,

4.3 $$\begin{aligned} &E.[\mathfrak{f}(\mathfrak{m}_{\partial D})^2] \\ &= 2E.\left[\int_0^{\mathfrak{m}_{\partial D}} [\mathfrak{f}(\mathfrak{m}_{\partial D}) - \mathfrak{f}(t)]\mathfrak{f}(dt)\right] \\ &= 0\,, \end{aligned}$$

and hence

4.4 $$\begin{aligned} 0 &= E.[\mathfrak{f}(\mathfrak{m}_{\partial D}) \mid \boldsymbol{B}_{t \wedge \mathfrak{m}_{\partial D}+}] = E.[\mathfrak{f}(\mathfrak{m}_{\partial D}) - \mathfrak{f}(t) \mid \boldsymbol{B}_{t \wedge \mathfrak{m}_{\partial D}+}] + \mathfrak{f}(t) \\ &= \mathfrak{f}(t) \qquad t \leq \mathfrak{m}_{\partial D} \end{aligned}$$

as desired.

It is a simple matter to deduce from this that *two additive functionals with the same associated measure are the same*; indeed, the difference p of two solutions of 3.2 satisfies $-p = \alpha \int Gpde$, which implies

4.5 $$0 \geq -\int p^2 de = \mathfrak{E}(pde) = \int Gpdepde\,,$$

and it follows that two additive functionals with the same associated measure have the same p_α, hence the same $\alpha^{-1}(1-p_\alpha) = E.[\mathfrak{f}_\alpha(\mathfrak{m}_{\partial D})]$ ($< \alpha^{-1} < +\infty$), and hence the same $\mathfrak{f}_\alpha$. But this means that the two functionals are the same up to time $\mathfrak{m}_{\partial D}$, and, to finish the proof, it is enough to make D swell out to R^d.

5. CONSTRUCTION OF AN ADDITIVE FUNCTIONAL FROM ITS ASSOCIATED MEASURE.

Given a *smooth* non-negative measure e, *our task is to find an additive functional* $\mathfrak{f}$ *with* e *as its associated measure.*

Consider, for this purpose, a non-negative measure e on a bounded open figure D with *bounded potential* $p=\int Gde$ and *finite energy* $\mathfrak{E}(e)=\int Gdede$, let p_n be the potential $\int Gf_n db$ of

5.1 $\quad f_n = n(p-E.[p(x(n^{-1})), n^{-1}<\mathfrak{m}_{\partial D}])$,

let

5.2 $\quad \mathfrak{f}_n(t) = \int_0^{t\wedge\mathfrak{m}_{\partial D}} f_n(x(s))ds$,

and let us construct a functional $\mathfrak{f}$ associated with e as a limit of $\mathfrak{f}_n$ with the aid of the following simple lemmas:

a) *p is a Brownian excessive function; esp.* $0\le f_n$.

b) $E.[p(x(t)), t<\mathfrak{m}_{\partial D}]\downarrow 0$ *inside D as* $t\uparrow+\infty$.

c) $p_n = n\int_0^{n^{-1}} E.[p(x(t)), t<\mathfrak{m}_{\partial D}]dt$ *increases to p inside D as* $n\uparrow +\infty$.

d) $\lim_{n\uparrow+\infty}\mathfrak{E}(e-f_n db)=0$, *where $\mathfrak{E}$ is the energy* $\mathfrak{E}(e)\equiv\int Gdede$.

e) $E.[\mathfrak{f}_n(+\infty)\,|\,\boldsymbol{B}_{t\wedge\mathfrak{m}_{\partial D}+}]\equiv\mathfrak{l}_n(t)$
$= p_n(x(t))+\mathfrak{f}_n(t) \qquad \mathfrak{m}_{\partial D}>t\ge 0$
$= \mathfrak{f}_n(\mathfrak{m}_{\partial D}) \qquad \mathfrak{m}_{\partial D}\le t$,

i.e., $\mathfrak{l}_n$ *is a martingate with respect to the fields* $\boldsymbol{B}_{t\wedge\mathfrak{m}_{\partial D}+}$; *moreover,* $\mathfrak{l}_n$ *is continuous in t.*

f) $P.[\max_{t\ge 0}|\mathfrak{l}_n(t,w_s^+)-\mathfrak{l}_m(t,w_s^+)|>\varepsilon]$
$\le \textit{constant}\times\varepsilon^{-2}s^{-d/2}\sqrt{\mathfrak{E}(f_n db-f_m db)}$.

g) $P.[\lim_{n\uparrow+\infty}\mathfrak{f}_n(t,w_s^+)=\mathfrak{f}(t,w_s^+), t\ge s>0]=1$.

where the limit is taken as $n\uparrow+\infty$ via suitable $n_1<n_2<etc.$, $\mathfrak{f}(t)$ *is continuous,* $\mathfrak{f}(0)\equiv\mathfrak{f}(0+)=0$, *and* $\mathfrak{f}(t)=\mathfrak{f}(s)+\mathfrak{f}(t-s,w_s^+)$ $(\mathfrak{m}_{\partial D}\ge t\ge s)$.

h) $E.[\mathfrak{f}(\mathfrak{m}_{\partial D})]=p$.

Because p is a potential, it is excessive; a) is obvious from this, b) is obvious from the bound

5.3 $E_\cdot[p(x(t)), t < \mathfrak{m}_{\partial D}] \leq \|p\|_\infty P_\cdot(t < \mathfrak{m}_{\partial D}) \downarrow 0 \qquad t \uparrow +\infty,$

and c) follows from b):

5.4 $$p_n = E_\cdot\left[\int_0^{\mathfrak{m}_{\partial D}} f_n(x(s))ds\right]$$
$$= n\int_0^{+\infty} dsE_\cdot[p(x(s)) - E_{x(s)}[p(x(n^{-1})), n^{-1} < m_{\partial D}], s < \mathfrak{m}_{\partial D}]$$
$$= n\int_0^{+\infty} ds[E_\cdot[p(x(s)), s < \mathfrak{m}_{\partial D}] - E_\cdot[p(x(s+n^{-1})), s+n^{-1} < \mathfrak{m}_{\partial D}]]$$
$$= n\int_0^{n^{-1}} E_\cdot[p(x(s)), s < \mathfrak{m}_{\partial D}]ds \uparrow p \qquad n \uparrow +\infty\,.$$

An application of c) establishes

5.5 $$\mathfrak{E}(e) = \int pde \geq \int p_n de = \int pf_n db \geq \mathfrak{E}(f_n db) \uparrow \mathfrak{E}(e) \qquad n \uparrow +\infty,$$

and this implies d):

5.6 $$\lim_{n\uparrow+\infty} \mathfrak{E}(e - f_n db) = \lim_{n\uparrow\infty} [\mathfrak{E}(e) - 2\int p_n de + \mathfrak{E}(f_n db)] = 0\,.$$

Because $t \wedge \mathfrak{m}_{\partial D}$ is a Markov time,

5.7 $$E_\cdot[\mathfrak{f}_n(\mathfrak{m}_{\partial D}) \mid B_{t\wedge\mathfrak{m}_{\partial D}+}] = E_\cdot\left[\int_{t\wedge\mathfrak{m}_{\partial D}}^{\mathfrak{m}_{\partial D}} f_n(x(s))ds \mid B_{t\wedge\mathfrak{m}_{\partial D}+}\right] + \mathfrak{f}_n(t)$$
$$= E_{x(t\wedge\mathfrak{m}_{\partial D})}\left[\int_0^{\mathfrak{m}_{\partial D}} f_n ds\right] + \mathfrak{f}_n(t) = p_n(x(t)) + \mathfrak{f}_n(t) \equiv \mathfrak{l}_n(t) \quad t < \mathfrak{m}_{\partial D},$$

i.e., $\mathfrak{l}_n$ is martingale, and since p_n is continuous and tends to 0 along the Brownian path ($t \uparrow \mathfrak{m}_{\partial D}$), $\mathfrak{l}_n$ is continuous. e) is now established, and f) follows from Doob's submartingale extension of Kolmogorov's inequality [4], the Schwarz inequality

5.8 $$\left(\int G de_1 de_2\right)^2 \leq \mathfrak{E}(e_1)\mathfrak{E}(e_2)$$

(see H. Cartan [2]), and the resulting

5.9 $$E_\cdot[|\mathfrak{l}_n(+\infty, w_s^+) - \mathfrak{l}_m(+\infty, w_s^+)|^2, s < \mathfrak{m}_{\partial D}]$$
$$\leq E_\cdot[E_{x(s)}[|\mathfrak{f}_n(\mathfrak{m}_{\partial D}) - \mathfrak{f}_m(\mathfrak{m}_{\partial D})|^2], s < \mathfrak{m}_{\partial D}]$$
$$\leq E_\cdot\left[\int G(x(s), b)(f_n - f_m)(p_n - p_m)db, s < \mathfrak{m}_{\partial D}\right]$$

$$= \iint G(a, b) e_{nm}(da) E_\bullet[G(x(s), b), s < \mathfrak{m}_{\partial D}] e_{nm}(db)$$

$$e_{nm}(db) = (f_n - f_m) db$$

$$\leq \mathfrak{E}(e_{nm})^{1/2} \mathfrak{E}(E_\bullet[G(x(s), b), s < \mathfrak{m}_{\partial D}] de_{nm})^{1/2}$$

$$\leq \mathfrak{E}(e_{nm})^{1/2} \textit{ constant } \times s^{-d/2} \mathfrak{E}(e_{nm})^{1/2}.$$

Choose $n_1 < n_2 <$ etc. so as to make $P_\bullet[\max_{t \geq 0} |\mathfrak{f}_n(t, w_s^+) - \mathfrak{f}_s(t)| \downarrow 0, s > 0] = 1$, where $\mathfrak{f}_s$ is a continuous function of t. Because $p = \lim_{n \uparrow \infty +} p_n$ is continuous and tends to 0 along the Brownian path $(t \uparrow \mathfrak{m}_{\partial D})$, it follows that $\mathfrak{f}_s(t) = \lim_{n \uparrow +\infty} \mathfrak{f}_n(t, w_s^+)$ is continuous $(t \geq 0)$ and additive $(t \leq \mathfrak{m}_{\partial D}(w_s^+))$; moreover

5.10 $\quad E_\bullet[\mathfrak{f}_n(+\infty)^2] = 2 \int G p_n f_n db \leq 2 \|p\|_\infty^2 < +\infty$

implies

5.11 $\quad E_\bullet[\mathfrak{f}_s(t-s)] = \lim_{n \uparrow +\infty} E_\bullet[\mathfrak{f}_n(+\infty, w_s^+) - \mathfrak{f}_n(+\infty, w_t^+)]$

$$= \lim_{n \uparrow +\infty} [E_\bullet(p_n(x(s)), s < \mathfrak{m}_{\partial D}) - E_\bullet(p_n(x(t)), t < \mathfrak{m}_{\partial D})]$$

$$= E_\bullet(p(x(s)), s < \mathfrak{m}_{\partial D}) - E_\bullet(p(x(t)), t < \mathfrak{m}_{\partial D}),$$

and, to finish the proof of g) and h), it is enough to define

5.12 $\quad \mathfrak{f}(t) = \lim_{s \downarrow 0} \mathfrak{f}_s(t)$

and to make $s \downarrow 0$ and $t \uparrow +\infty$ in 5.11.

Now take a smooth measure e, choose $B_n \uparrow D$ as needed for 1.8 with the additional property that $e|B_n$ has *finite energy* $\mathfrak{E}(e|B_n) = \int_{B_n \times B_n} G de de$[6], let $\mathfrak{f}_n$ be the additive functional associated with $e|B_n$ as in 5.12 above, and note that

5.13 $\quad E_\bullet\left[\int_0^{\mathfrak{m}_{\partial D}} f d\mathfrak{f}_n\right] = \int_{B_n} G f de \qquad f \geq 0$

as in e) of section 3. It follows that if f is the indicator function

---

[6] $\mathfrak{E}(e|B_n) < +\infty$ is achieved as follows: take $\dot{D} \supset D$ with $\partial\dot{D}$ at a positive distance from ∂D, choose $\dot{B}_n \uparrow \dot{D}$ as needed for 1.8, and let $B_n = \dot{B}_n \cap D$; then $\mathfrak{E}(e|B_n) \leq \int_{B_n} de \int_{B_n} \dot{G} de \leq n e(B_n) \leq n (\inf_{D \times D} \dot{G})^{-1} \sup_D \int_{B_n} \dot{G} de \leq n^2 (\inf_{D \times D} \dot{G})^{-1} < +\infty$.

of B_m, then the functionals $\int_0^t f d\mathfrak{f}_n$ $(n>m)$ and $\int_0^t f d\mathfrak{f}_m$ have the same (bounded) mean:

5.14 $$E.\left[\int_0^{\mathfrak{m}_{\partial D}} f d\mathfrak{f}_n\right] = E.\left[\int_0^{\mathfrak{m}_{\partial D}} f d\mathfrak{f}_m\right] = \int_{Bm} G de$$

and are therefore identical up to time $\mathfrak{m}_{\partial D}$ according to the first uniqueness lemma for additive functionals of section 4. But this means that $\mathfrak{f}_n = \mathfrak{f}_m$ up to the exit time from B_m, and since this exit time $=\mathfrak{m}_{\partial D}$ for large m, it is legitimate to define a functional $\mathfrak{f}$ for $t \leq \mathfrak{m}_{\partial D}$ by means of

5.15 $$\mathfrak{f}(t) = \mathfrak{f}_n(t) \qquad t \leq \inf(t: x(t) \notin B_n),\ n \geq 1\,.$$

Introducing $p_\alpha = E.[e^{-\alpha \mathfrak{f}(\mathfrak{m}_{\partial D})}]$ and using the method of 3.5, 5.13 implies

5.16 $$1 - p_\alpha = \lim_{n \uparrow +\infty} \alpha E.\left[\int_0^{\mathfrak{m}_{\partial D}} e^{-\alpha[\mathfrak{f}(\mathfrak{m}_{\partial D}) - \mathfrak{f}(t)]} \mathfrak{f}_n(dt)\right]$$
$$= \lim_{n \uparrow +\infty} \alpha E.\left[\int_0^{\mathfrak{m}_{\partial D}} p_\alpha d\mathfrak{f}_n\right] = \lim_{n \uparrow +\infty} \alpha \int_{B_n} G p_\alpha de = \alpha \int G p_\alpha de\,;$$

in brief, e is the measure associated with $\mathfrak{f}$ as in 3.2.

Because two additive functionals with same associated measure are the same, it follows that if e is smooth and if $D_1 \subset D_2 \subset$ *etc.* swell out to R^d, then the functionals $\mathfrak{f}_n$ associated with $e_n = e \mid D_n$ as in 5.15 above determine an additive functional

5.17 $$\mathfrak{f}(t) = \mathfrak{f}_n(t) \qquad t < \mathfrak{m}_{\partial D_n},\ n \geq 1$$

depending upon e alone, for which 3.2 holds for each D.

$\mathfrak{f}$ is additive for *almost all* Brownian paths, but of course it can be modified on a negligeable class of paths so as to be Borel in the pair (t, w) and to satisfy 1.1, 1.2, 1.3, 1.4, as *identities*; with *this modification, $\mathfrak{f}$ is the additive functional associated with e.*

6. DIFFUSIONS WITH BROWNIAN HITTING PROBABILITIES.

To avoid confusion, let w, x, $\mathfrak{m}$, B, $\boldsymbol{B}$, $\boldsymbol{B}_{\mathfrak{m}+}$, *etc.* be used to describe the Brownian motion, introduce the same paths, times, events, and fields with the new names $\dot{w}$, $\dot{x}$, $\dot{\mathfrak{m}}$, $\dot{B}$, $\dot{\boldsymbol{B}}$, $\dot{\boldsymbol{B}}_{\dot{\mathfrak{m}}+}$, *etc.*, and take a *new motion* $\dot{\boldsymbol{D}}$ with probabilities $\dot{P}_a(\dot{B})$.

$\dot{D}$ is said to be a *diffusion* if it *starts afresh* at each Markov time $\dot{\mathfrak{m}}$:

6.1 $\dot{P}_a[\dot{w}^+_{\dot{\mathfrak{m}}} \in \dot{B} \mid \dot{\boldsymbol{B}}_{\dot{\mathfrak{m}}+}] = \dot{P}_b(\dot{B}) \qquad a \in E^d,\ \dot{B} \in \dot{\boldsymbol{B}},\ b \equiv \dot{x}(\dot{\mathfrak{m}})$;

it is said to have Brownian *hitting probabilities* if, for each bounded open D,

6.2 $\dot{P}_a[\dot{x}(\dot{\mathfrak{m}}_{\partial D}) \in db] = P_a[x(\mathfrak{m}_{\partial D}) \in db] = h_{\partial D}(a, db)$
$a \in D,\ db \subset \partial D$.

Given such a diffusion with Brownian hitting probabilities

6.3a $\dot{p}_\alpha \equiv \dot{E}.[e^{-\alpha \dot{\mathfrak{m}}_{\partial D}}]$ [7]

solves

6.3b $1 - \dot{p}_\alpha = \alpha \int G \dot{p}_\alpha de$,

where $e \geq 0$ is independent of α and of D. e is the so-called *speed measure* of the diffusion; the speed measure of the Brownian motion is the Lebesgue measure db. Because of 6.2, $\dot{P}.[\dot{\mathfrak{m}}_{\partial D} < +\infty] = 1$ and $\dot{p}_\alpha > 0$, which implies that *e is smooth, i.e., that it is the associated measure of some additive functional* $\mathfrak{f}$ of the Brownian path; moreover, *e is positive on the neighborhoods of H. Cartan's fine topology, and this implies that its associated additive functional satisfies*

6.4 $P.[\mathfrak{f}(s) < \mathfrak{f}(t),\ 0 \leq s < t] = 1$.

Introducing the *inverse function* $\mathfrak{f}^{-1}$ of $\mathfrak{f}$, it turns out that

6.5 $\dot{P}.(\dot{B}) = P.[x(\mathfrak{f}^{-1}) \in \dot{B}] \qquad \dot{B} \in \dot{\boldsymbol{B}}$;

in brief, $\dot{D}$ is identical in law to the Brownian motion run with the stochastic clock $\mathfrak{f}^{-1}$; moreover, *each smooth measure e, positive on fine neighborhoods, is the speed measure of a diffusion with Brownian hits.*

The proofs are carried out in the next few sections.

---

[7] $\dot{E}.(f) = \int f d\dot{P}.$.

7. SPEED MEASURES

Beginning with the speed measure e of 6.3b, if a $a \in R^d$ and if $\dot{\mathfrak{m}}_\varepsilon = \min(t : |\dot{x}(t) - a| \geq \varepsilon)$, then $\dot{\mathfrak{m}}_{0+} = \lim_{\varepsilon \downarrow 0} \dot{\mathfrak{m}}_\varepsilon$ satisfies

$$7.1 \qquad \dot{E}_a(e^{-\dot{\mathfrak{m}}_{0+}}) = \dot{E}_a[e^{-\dot{\mathfrak{m}} - \dot{\mathfrak{m}}(w^+_{\dot{\mathfrak{m}}})}] = \dot{E}_a(e^{-\dot{\mathfrak{m}}_{0+}})^2, \qquad \dot{\mathfrak{m}} = \dot{\mathfrak{m}}_{0+}$$

and since $\dot{E}_a(e^{-\dot{\mathfrak{m}}_{0+}}) = 0$ implies $\dot{P}_a[\dot{\mathfrak{m}}_{0+} = +\infty] = 1$, violating the fact that the dot motion has Brownian hitting probabilities, it follows that $\dot{P}_a[\dot{\mathfrak{m}}_{0+} = 0] = 1$.

Because

$$7.2 \qquad 1 - \dot{p}_\alpha = \alpha \dot{E}.\left[\int_0^{\dot{\mathfrak{m}}_{\partial D}} e^{-\alpha(\dot{\mathfrak{m}}_{\partial D} - t)} dt\right] = \alpha \dot{E}.\left[\int_0^{\dot{\mathfrak{m}}_{\partial D}} \dot{p}_\alpha(x(t)) dt\right],$$

$1 - \dot{p}_\alpha$ satisfies

$$7.3 \qquad 1 - \dot{p}_\alpha(a) \geq \alpha \dot{E}_a\left[\int_{\dot{\mathfrak{m}}_{\partial \dot{D}}}^{\dot{\mathfrak{m}}_{\partial D}} \dot{p}_\alpha dt\right] = \int h_{\partial \dot{D}}(a, db)(1 - \dot{p}_\alpha) \qquad a \in \dot{D} \subset D,$$

and, using 7.3 and $\dot{P}_a[\dot{\mathfrak{m}}_{0+} = 0] = 1$ to establish

$$7.4\text{a} \qquad \int h_{\partial \dot{D}}(a, db)(1 - \dot{p}_\alpha) \downarrow 0 \qquad \dot{D} \uparrow D$$

$$7.4\text{b} \qquad \int h_{|b-a|=\varepsilon}(a, db)(1 - \dot{p}_\alpha) \uparrow 1 - \dot{p}_\alpha \qquad \varepsilon \downarrow 0,$$

it appears that $1 - \dot{p}_\alpha$ is the potential $\alpha \int G de_\alpha$ of a non-negative charge distribution αe_α

It remains to verify that $\dot{p}_\alpha^{-1} de_\alpha \equiv de$ is independent of α and of D, which is done with the aid of the additive functional $\dot{\mathfrak{f}}_\alpha(t) = \int_0^t \dot{p}_\alpha(\dot{x}(s)) ds$ $(t < \dot{\mathfrak{m}}_{\partial D})$ and the method of section 3; in outline,

$$7.5 \qquad 1 - \dot{p}_\alpha = \alpha \dot{E}_\alpha[\dot{\mathfrak{f}}_\alpha(\dot{\mathfrak{m}}_{\partial D})] = \alpha \int G de_\alpha$$

implies

$$7.6 \qquad \dot{E}\left[\int_0^{\dot{\mathfrak{m}}_{\partial D}} f(\dot{x}(s)) \dot{\mathfrak{f}}_\alpha(ds)\right] = \int G f de_\alpha$$

as in e) of section 3, and 3.9 follows as before, *etc.*.

8. TWO DIFFUSIONS WITH BROWNIAN HITS AND THE SAME SPEED MEASURE ARE THE SAME.

Consider a pair of diffusions with Brownian hitting probabilities and the *same* speed measure.

Because 6.3b=3.2 has unique solutions as noted in section 4, both diffusions have the same $\dot{p}_\alpha = \dot{E}_\cdot[e^{-\alpha\dot{\mathfrak{m}}_{\partial D}}]$ and hence the same

8.1 $$\dot{G}_{0+} : f \to \dot{E}_\cdot\left[\int_0^{\dot{\mathfrak{m}}_{\partial D}} f(\dot{x})\dot{\mathfrak{f}}_\varepsilon(dt)\right] = \int G f \dot{p}_\varepsilon de$$

$$\dot{\mathfrak{f}}_\varepsilon(t) \equiv \int_0^t \dot{p}_\varepsilon(\dot{x}(s))ds \qquad t \leq \dot{\mathfrak{m}}_{\partial D}.$$

Choose $\varepsilon > 0$ and introduce the Green operators

8.2 $$\dot{G}_\alpha : f \to \dot{E}_\cdot\left[\int_0^{\dot{\mathfrak{m}}_{\partial D}} e^{-\alpha\dot{\mathfrak{f}}_\varepsilon(t)} f(\dot{x})\dot{\mathfrak{f}}_\varepsilon(dt)\right] \qquad \alpha > 0;$$

then

8.3 $$\begin{aligned} \alpha\dot{G}_{0+}\dot{G}_\alpha f &= \alpha\dot{E}_\cdot\left[\int_0^{\dot{\mathfrak{m}}_{\partial D}} \dot{\mathfrak{f}}_\varepsilon(dt)(\dot{G}_\alpha f)(\dot{x})\right] \\ &= \alpha\dot{E}_\cdot\left[\int_0^{\dot{\mathfrak{m}}_{\partial D}} \dot{\mathfrak{f}}_\varepsilon(dt)\int_t^{\dot{\mathfrak{m}}_{\partial D}} e^{-\alpha[\dot{\mathfrak{f}}_\varepsilon(s)-\dot{\mathfrak{f}}_\varepsilon(t)]} f(\dot{x}(s))\dot{\mathfrak{f}}_\varepsilon(ds)\right] \\ &= \alpha\dot{E}_\cdot\left[\int_0^{\dot{\mathfrak{m}}_{\partial D}} e^{-\alpha\dot{\mathfrak{f}}_\varepsilon(s)} f(\dot{x})\dot{\mathfrak{f}}_\varepsilon(ds)\int_0^s e^{\alpha\dot{\mathfrak{f}}_\varepsilon} d\dot{\mathfrak{f}}_\varepsilon\right] \\ &= \dot{E}_\cdot\left[\int_0^{\dot{\mathfrak{m}}_{\partial D}} (1-e^{-\alpha\dot{\mathfrak{f}}_\varepsilon}) f d\dot{\mathfrak{f}}_\varepsilon\right] = \dot{G}_{0+}f - \dot{G}_\alpha f, \end{aligned}$$

i.e.,

8.4 $$\dot{G}_\alpha = \dot{G}_{0+} - \alpha\dot{G}_{0+}\dot{G}_\alpha,$$

and, using the bound

8.5 $$\dot{G}_{0+}1 = \int G\dot{p}_\varepsilon de = \varepsilon^{-1}(1-\dot{p}_\varepsilon) \leq \varepsilon^{-1} < +\infty,$$

the obvious iteration of 8.4 implies

8.6 $$\dot{G}_\alpha = \sum_{n\geq 0}(-)^n\alpha^n\dot{G}_{0+}^{n+1} \qquad \alpha < \varepsilon.$$

Because $\dot{G}_\alpha$ is a Laplace transform in its dependence on α, 8.6 implies that both diffusions have the same Green operators ($\alpha \geq 0$) and hence the same

8.7 $$\dot{E}_\cdot\left[\int_0^{\dot{\mathfrak{m}}_\infty} e^{-\alpha t} f(\dot{x})dt\right] = \lim_{D\uparrow R^d}\lim_{\varepsilon\downarrow 0}\dot{G}_\alpha f.$$

But this implies that both diffusions have the same

8.8 $$\dot{P}_a[\dot{x}(t) \in db,\ t < \dot{\mathfrak{m}}_\infty] \qquad (t, a, b) \in (0, +\infty)\times R^{2d}$$

and hence are the same in all respects as stated in the section title.

9. SPEED MEASURES ARE POSITIVE ON FINE NEIGHBORHOODS.

Given a point a of bounded open D and a (Brownian) excessive function p on D, the set of points $b \in D$ at which $|p(b)-p(a)|<n^{-1}$ is said to be a *fine neighborhood* of a; the corresponding topology is called the *fine topology of H. Cartan* [2]. E. B. Dynkin [7] has pointed out that a point a is in the fine interior of $B \subset R^d$ if and only if almost all Brownian paths starting at a remain in B for some positive time. Because an excessive function in 1-dimension is concave and hence continuous, the 1-dimensional *fine* topology is the same as the usual one; in higher dimensions it is different.

Given a diffusion with Brownian hitting probabilities, *its speed measure e has to be positive on fine neighborhoods*; for the proof, it is enough to verify that if $Z \subset R^d$ is bounded and Borel and if $e(Z)=0$, then Z has no fine interior.

Choose an open ball D and a decreasing series of open figures D_n so as to have

9.1a $\quad D \supset D_1 \supset D_2 \supset etc. \supset Z$

9.1b $\quad \int_{D_n} G_D \dot{p}_1 de \downarrow 0 \qquad n \uparrow +\infty, \ \dot{p}_1 = \dot{E}_\cdot[e^{-\mathfrak{m}_{\partial D}}]$,

and let

9.2 $\quad \dot{\mathfrak{f}}_1(t) = \int_0^t \dot{p}_1(\dot{x}(s))ds \qquad t \leq \mathfrak{m}_{\partial D}$.

Because of

9.3a $\quad 1-\dot{E}_a[e^{-\dot{\mathfrak{f}}_1(\partial\mathfrak{m}_{D_n})}] = \int_{D_n} G_{D_n}\dot{p}_1 de \leq \int_{D_n} G_D \dot{p}_1 de \downarrow 0$

$$n \uparrow +\infty, \ a \in Z$$

9.3b $\quad \dot{p}_1 > 0$,

it is apparent that

9.4a $\quad 1 = \dot{P}_a[\mathfrak{m}_{\partial D_n} \downarrow 0, n \uparrow +\infty]$

$\qquad = \dot{P}_a[\lim_{n \uparrow +\infty} \dot{x}(\mathfrak{m}_{\partial D_n}) = a]$

$\qquad = \lim_{n \uparrow +\infty} \int h_{\partial D_n}(a, db) e^{-|b-a|} \qquad a \in Z$.

But this final expression depends upon the (Brownian) hitting probabilities $h_{\partial D}$ alone; thus,

9.4b $\quad P_a[\lim_{n\uparrow+\infty} x(\mathfrak{m}_{\partial D}) = a] = 1 \qquad a \in Z$

for the Brownian motion also, and since almost no Brownian path meets its starting point at a *positive* time, it follows that

9.5 $\quad P_a[\mathfrak{m}_{\partial D_n} \downarrow 0,\ n \uparrow +\infty] = 1 \qquad a \in Z;$

an application of Dynkin's description of fine neighborhoods completes the proof.

10. SPEED MEASURES GIVE RISE TO INCREASING ADDITIVE FUNCTIONALS.

Because the speed measure of a diffusion with Brownian hits is smooth, it has associated with it an additive functional $\mathfrak{f}$ of the Brownian path; moreover, e is positive on fine neighborhoods, and this is reflected in the fact that $\mathfrak{f}$ *is increasing as in* 6.4: $\mathfrak{f}(s) < \mathfrak{f}(t) (s < t)$.

It will be enough to verify that the set A of points a at which

10.1 $\quad P_a(\mathfrak{m} > 0) = P_a[\mathfrak{f}(t) = 0 \text{ for some } t > 0] = 1$
$\qquad \mathfrak{m} = \inf(t : f(t) > 0)$

is vaccuous; note that A is Borel and that $P_{\cdot}(\mathfrak{m} > 0) = 0$ or 1 according to Blumenthal's 01 law.

Because A is either *void* or *fine open* and e is positive on fine neighborhoods, it is enough to show that $e(A) = 0$, and, for this, it is enough to show that for each bounded open D, the points of D at which $P_{\cdot}[\mathfrak{f}(\mathfrak{m}_{\partial D}) > 0] < 1$ have e-mass 0. But this is immediate on letting $\alpha \uparrow +\infty$ in

10.2 $\quad \alpha^{-1}(1 - p_\alpha) = \int G p_\alpha de \qquad p_\alpha = E_{\cdot}[e^{-\alpha \mathfrak{f}(\mathfrak{m}_{\partial D})}].$

11. WHEN IS $\mathfrak{f}(+\infty) = +\infty$?

In making up the stochastic clock $\mathfrak{f}^{-1}$ for use in 6.5, two cases arise according as $\mathfrak{f}(+\infty) = +\infty$ or not, and it is desirable to have a test for this.

Because $p \equiv P_{\cdot}[\mathfrak{f}(+\infty) < +\infty]$ satisfies

11.1 $p(a) = P_a[\mathfrak{f}(\mathfrak{m}_{\partial D}) < +\infty,\ \mathfrak{f}(+\infty, w^+_{\mathfrak{m}_{\partial D}}) < +\infty]$

$$= E_a[p(x(\mathfrak{m}_{\partial D}))] = \int h_{\partial D}(a, db)p(b) \qquad a \in D$$

for bounded open D, it is harmonic and since $p \geq 0$, it must be constant; moreover, letting first t and then $n \uparrow +\infty$ in

11.2 $P_\cdot[\mathfrak{f}(+\infty) < n] \leq E_\cdot[\mathfrak{f}(t) < n,\ p(x(t))] \leq P_\cdot[\mathfrak{f}(t) < n]p$,

it appears that $p \leq p^2$ and hence that $p = 0$ or 1.

Our test states that $p = 1$ *if and only if one of the following conditions is met*:

a) $d \geq 3$ *and* $1 - p_1 = \int G p_1 de$ *admits a solution* $0 < p_1 \leq 1$ *on the whole of* R^d, *where* G *is the function of* R^d.

b) $\mathfrak{G} p_1 = p_1$ *admits a solution* $0 < p_1 \leq 1$ *on the whole of* R^d, *where* $\mathfrak{G} p_1$ *is the negative of the Radon-Nikodym derivative of the Riesz measure of* p_1 *with respect to* e (see section 13 for the meaning of $\mathfrak{G}$).

c) $d \geq 3$ *and* $R^d = A \cup B$, *where* A *is thin at* ∞ *in the sense that* $P_\cdot[x(t) \in A$ *for some* $t > n] \downarrow 0$ $(n \uparrow \infty)$, *and* $\int_B G de < +\infty$.[8]

Beginning with $p = 1$, $p_1 \equiv E_\cdot[e^{-\mathfrak{f}(+\infty)}]$ is positive and ≤ 1, and, since $G_D \uparrow G$ as $D \uparrow R^d$,

11.3 $1 - p_1 = \lim_{D \uparrow R^d} E_\cdot[e^{-\mathfrak{f}(+\infty,\, w^+_{\mathfrak{m}_{\partial D}})}] - p_1$

$$= \lim_{D \uparrow R^d} \int G_D p_1 de = \int G p_1 de.$$

Because $G \equiv +\infty$ in case $d = 2$, it follows that $p = 1$ implies a), that a) implies b) is clear, that c) implies $p = 1$ is evident from

11.4 $E_\cdot\left[\int_0^{+\infty} f d\mathfrak{f}\right] = \int_B G de < +\infty$,

where f is the indicator function of B, and now it remains to verify that b) implies c).

But, if $0 < p_1 \leq 1$ is a solution of $\mathfrak{G} p_1 = p_1$ and if h_1 is its

---

[8] $A \subset R^d (d \geq 3)$ is thin at ∞ if and only if (Wiener's test) $\sum_{n \geq 1} 2^{-n(d-2)} C(A_n) < +\infty$, where A_n is the meet of A with the spherical $2^{n-1} \leq |b| < 2^n$ and C is the d-dimensional Newtonian capacity; for the proof in the case of the d-dimensional random walk, see K. Itô and H. P. McKean, Jr. [12].

harmonic part $E_{\bullet}[p_1(x(\mathfrak{m}_{\partial D}))]$ inside D, then $h_1 - p_1 = \int G_D p_1 de$ inside D, and, as $D \uparrow R^d$, h_1 decreases to a non-negative (and hence constant) harmonic function $p_1(\infty)$ such that $p_1(\infty) - p_1 = \int G p_1 de$. R^d is now split into $A = (p_1 < \frac{1}{2} p_1(\infty))$ and $B = (p_1 \geq \frac{1}{2} p_1(\infty))$ and the fact that $p_1(\infty) - p_1$ excessive is used to ensure that p_1 has a limit along the Brownian path as $t \uparrow +\infty$, permitting us to conclude from

11.4a $\quad p_1 \leq p_1(\infty)$

11.4b $\quad p_1(\infty) = \lim_{D \uparrow R^d} E_{\bullet}[p_1(x(\mathfrak{m}_{\partial D}))] = E_{\bullet}[\lim_{t \uparrow +\infty} p_1(x(t))]$

that A is thin at ∞. Because $0 < p_1(\infty)$,

11.5 $\quad \int_B G de \leq 2 p_1(\infty)^{-1} \int G p_1 de < +\infty$,

and this completes the verification of c).

12. PERFORMING THE TIME SUBSTITUTION.

Coming to the actual time substitution $t \to \mathfrak{f}^{-1}$ which is supposed to send the Brownian motion into the diffusion $\dot{D}$, let $\mathfrak{f}$ be modified on a negligeable class of Brownian paths so as to have

12.1 $\quad 0 = \mathfrak{f}(0) \leq \mathfrak{f}(t \pm) = \mathfrak{f}(t) < +\infty$

12.2 $\quad \mathfrak{f}(t) = \mathfrak{f}(s) + \mathfrak{f}(t-s, w_s^+) \qquad t \geq s$

12.3 $\quad \mathfrak{f}(t) > \mathfrak{f}(s) \qquad t > s$

as *identities*, let x^{-1} denote the sample path

12.4 $\quad w^{-1}: t \to x^{-1}(t) \equiv x^{-1}(t, w^{-1}) \equiv x[\mathfrak{f}^{-1}(t, w), w]$,

note that this path is continuous even if $\mathfrak{f}(+\infty) < +\infty (d \geq 3)$, and let us check that *the motion* $\dot{D}$ *with sample paths*

12.5a $\quad \dot{w}: t \to \dot{x}(t)$

and probabilities

12.5b $\quad \dot{P}_a(\dot{B}) \equiv P_a(w^{-1} \in \dot{B})$

is the diffusion with Brownian hitting probabilities and speed measure e.

Beginning with the proof of the *diffusive character* 6.1 of this motion, the problem is to check that if $\dot{\mathfrak{m}}$ is a Markov time and if $\dot{B} \in \dot{\boldsymbol{B}}$, then

12. 6 $\dot{P}_a[\dot{w}^+_{\dot{\mathfrak{m}}} \in \dot{B} \mid \dot{\boldsymbol{B}}_{\dot{\mathfrak{m}}+}] = \dot{P}_b(\dot{B}) \qquad b \equiv \dot{x}(\dot{\mathfrak{m}})$.

Given such a Markov time $\dot{\mathfrak{m}}$,

12. 7 $\mathfrak{m}(w) \equiv \mathfrak{f}^{-1}(\dot{\mathfrak{m}}(w^{-1}), w)$

is a Markov time for the Brownian path ; indeed, using Galmarino's test, if

12. 8a $\mathfrak{m}(u) < t$

12. 8b $x(s, u) = x(s, v) \qquad s \leq t$,

then

12. 9a $\mathfrak{f}(s, u) = \mathfrak{f}(s, v) \qquad s \leq t$

12. 9b $\mathfrak{f}^{-1}(s, u) = \mathfrak{f}^{-1}(s, v) \leq t \qquad s \leq \dot{t} = \mathfrak{f}(t, u) = \mathfrak{f}(t, v)$,

and it follows that

12. 10a $\dot{\mathfrak{m}}(u^{-1}) = \mathfrak{f}(\mathfrak{m}(u), u) < \mathfrak{f}(t, u) = \dot{t}$

12. 10b $x^{-1}(s, u^{-1}) = x[\mathfrak{f}^{-1}(s, u), u] = x[\mathfrak{f}^{-1}(s, v), v]$
$= x^{-1}(s, v^{-1}) \qquad s \leq \dot{t}$.

Because $\dot{\mathfrak{m}}$ was a Markov time for the dot motion, 12. 10 implies

12. 11a $\dot{\mathfrak{m}}(u^{-1}) = \dot{\mathfrak{m}}(v^{-1}) < \dot{t}$,

and an application of 12. 9b implies

12. 11b $\dot{\mathfrak{m}}(u) = \mathfrak{f}^{-1}(\dot{\mathfrak{m}}(u^{-1}), u) = \mathfrak{f}^{-1}(\mathfrak{m}(v^{-1}), v) = \mathfrak{m}(v)$;

in brief, 12. 8 *implies*, 12. 11b, as needed to conclude by Galmarino's test that $\mathfrak{m}$ is a Markov time.

Given $\dot{A} \in \dot{\boldsymbol{B}}_{\dot{\mathfrak{m}}+}$, if $A = (w : w^{-1} \in \dot{A})$ and if

12. 12a $\mathfrak{m}(u) < t$

12. 12b $x(s, u) = x(s, v) \qquad s \leq t$

12. 12c $u \in A$,

then, using 12. 10,

12. 13a $\dot{\mathfrak{m}}(u^{-1}) < \dot{t}$

12. 13b $x^{-1}(s, u^{-1}) = x^{-1}(x, v^{-1}) \qquad s \leq \dot{t}$

12. 13c $u^{-1} \in A$,

and it follows that $v^{-1}\in\dot{A}$, or, what is the same, that $v\in A$; thus, by Galmarino's test, $A\in \boldsymbol{B}_{\mathfrak{m}+}$, and now it appears that

12.14
$$\begin{aligned}
&\dot{P}_a[\dot{A},\ \dot{w}^+_{\dot{\mathfrak{m}}}\in\dot{B}]\\
&= P_a[w^{-1}\in\dot{A},\ (w^+_{\mathfrak{m}})^{-1}\in\dot{B}]\\
&= P_a[w\in A,\ (w^+_{\mathfrak{m}})^{-1}\in\dot{B}]\\
&= E_a[A,\ P_b(w^{-1}\in\dot{B})] \qquad b\equiv x(\mathfrak{m}) = x^{-1}(\dot{\mathfrak{m}}(w^{-1}),\ w^{-1})\\
&= E_a[w^{-1}\in\dot{A},\ \dot{P}_b(\dot{B})]\\
&= \dot{E}_a[\dot{A},\ \dot{P}_b(\dot{B})] \qquad b=\dot{x}(\dot{\mathfrak{m}}),
\end{aligned}$$

completing the proof of 12.6.

$\dot{D}$ is now identified as a *diffusion*; that it has *Brownian hitting probabilities* is clear, and to complete the discussion, it suffices to verify that *it has e as its speed measure.* But this is clear because

12.15 $$\mathfrak{m}^{-1}_{\partial D}\equiv\min(t: x^{-1}(t)\in\partial D) = \mathfrak{f}(\mathfrak{m}_{\partial D}),$$

and $\mathfrak{f}$ has e as its associated measure.

13. GENERATORS.

Given a diffusion with Brownian hitting probabilities, the *Green operators*

13.1 $$\dot{G}_\alpha: f\to\dot{E}_\cdot\left[\int_0^{+\infty}e^{-\alpha t}f(\dot{x})dt\right] \qquad \alpha>0$$

map into itself the space $\dot{C}(E^d)$ of real, bounded, fine-continuous functions having ordinary limits at ∞ in case $d\geq 3$; in fact, if $d\geq 3$, then $\dot{P}_\cdot[\min_{t\geq 0}|\dot{x}(t)|\geq n] = P_\cdot[\min_{t\geq 0}|x(t)|\geq n]$ tends to 1 at ∞, so that $\dot{G}_\alpha f$ tends to $\alpha^{-1}f(\infty)$, and, if $a\in R^d$ $(d\geq 2)$ and if the ball $D\ni a$ is so small that $1-\dot{p}_\alpha(a) = 1-\dot{E}_\cdot[e^{-\alpha\dot{\mathfrak{m}}_{\partial D}}]<n^{-1}$, then, inside the *fine neighborhood* $B=D\cap(\dot{p}_\alpha>1-n^{-1})$, the difference between

13.2a
$$\begin{aligned}
u &= \dot{G}_\alpha f\\
&= \dot{E}_\cdot\left[\int_0^{\dot{\mathfrak{m}}_{\partial D}}e^{-\alpha t}f(\dot{x})dt + e^{-\alpha\dot{\mathfrak{m}}_{\partial D}}\int_0^{+\infty}e^{-\alpha t}f(\dot{x}(t,\ \dot{w}^+_{\dot{\mathfrak{m}}_{\partial D}}))dt\right]\\
&= \dot{E}_\cdot\left[\int_0^{\dot{\mathfrak{m}}_{\partial D}}e^{-\alpha t}f(\dot{x})dt + e^{-\alpha\dot{\mathfrak{m}}_{\partial D}}u(\dot{x}(\dot{\mathfrak{m}}_{\partial D}))\right]
\end{aligned}$$

and the harmonic (and hence continuous) function

13.2b $\quad h = \dot{E}_{\bullet}[u(\dot{x}(\dot{\mathfrak{m}}_{\partial D}))]$

is not greater than

13.3 $\quad constant \times (1-\dot{p}_\alpha) < constant \times n^{-1}$.

Because

13.4 $\quad \dot{G}_\alpha - \dot{G}_\beta + (\alpha-\beta)\dot{G}_\alpha\dot{G}_\beta = 0 \qquad \alpha, \beta > 0$,

it is evident that $\dot{G}_\alpha$ maps our space of fine continuous functions onto some subspace $D(\dot{\mathfrak{G}})$ independent of α and that its null-space $\dot{G}_\alpha^{-1}(0)$ is likewise independent of α. But, for fine-continuous $f \in \dot{G}_\alpha^{-1}(0)$,

13.5 $$0 = \lim_{\beta\uparrow+\infty} \beta\dot{G}_\beta f = \dot{E}_{\bullet}\left[\lim_{\beta\uparrow+\infty} \beta\int_0^{+\infty} e^{-\beta t} f(\dot{x})\,dt\right] = f$$

according to E. B. Dynkin's description of fine neighborhoods; thus the null-space is trivial, $\dot{G}_\alpha$ is invertable, and another application of 13.4 verifies that $\dot{\mathfrak{G}} \equiv \alpha - \dot{G}_\alpha^{-1}$ acting on $D(\dot{\mathfrak{G}})$ is independent of α.

$\dot{\mathfrak{G}}$ is the so-called generator; it is *closed* in the sense that if $u_n \in D(\dot{\mathfrak{G}})$ and $\dot{\mathfrak{G}}u_n = f_n$ converge pointwise under fixed bounds to u and $f \in \dot{C}(E^d)$, then $u \in D(\dot{\mathfrak{G}})$ and $\dot{\mathfrak{G}}u = f$.

Consider the *differential operator*

13.6 $$\mathfrak{D}u = \frac{-e^u(db)}{e(db)} \qquad |b| < +\infty$$
$$= 0 \qquad b = \infty, \ d \geq 3$$

acting on the class $D(\mathfrak{D})$ of functions $u \in \dot{C}(E^d)$ such that

a) *inside each D, u is the sum of the harmonic function* $h = \int h_{\partial D}(a, db)u(b)$ *and the potential* $\int G_D de^u$ *of its Riesz measure* e^u.

b) $\int G_D |de^u|$ *is bounded.*

c) $\mathfrak{D}u$, *as described in* 13.6, *exists and belongs to* $\dot{C}(E^d)$.

$\dot{\mathfrak{G}}$ *is the closure* $\overline{\mathfrak{D}}$ *of* $\mathfrak{D}$ *in the topology of bounded pointwise convergence* as will now be explained.

Choose a fine-continuous function $0 < p_n \leq 1$ tending to 0 at ∞ in the ordinary topology such that $\int G_D p_n de$ is bounded for each bounded open D and $p_n \uparrow 1$ as $n \uparrow +\infty$, and introduce the additive functional $\dot{\mathfrak{f}}_n = \int_0^t p_n(\dot{x})ds \ (t \geq 0)$.

$\dot{x}(\dot{\mathfrak{f}}_n^{-1})$ *is a diffusion with Brownian hitting probabilities and speed measure* $de_n = p_n \times de$,

13. 7 $\dot{E}_\cdot[\min(t: \dot{x}(\dot{\mathfrak{f}}_n^{-1}) \in \partial D)] = \dot{E}_\cdot[\dot{\mathfrak{f}}_n(\dot{\mathfrak{m}}_{\partial D})] = \int G_D p_n de < +\infty$,

and it follows from 13. 2a that if u is in the domain of its generator $\dot{\mathfrak{G}}_n$, then

13. 8 $$u = -\dot{E}_\cdot\left[\int_0^{\dot{\mathfrak{m}}_{\partial D}} (\dot{\mathfrak{G}}_n u)(\dot{x}(t))\dot{\mathfrak{f}}_n(dt)\right] + E_\cdot[u(\dot{x}\mathfrak{m}_{\partial D}))]$$
$$= -\int G_D(\dot{\mathfrak{G}}_n u)p_n de + a \textit{ harmonic function},$$

which is a special case of a formula of E. B. Dynkin [6]; *in brief*,

13. 9 $$-e^u(db) = (\dot{\mathfrak{G}}_n u)p_n de \qquad u \in D(\dot{\mathfrak{G}}_n).$$

Choose $u = \dot{G}_1 f \in D(\dot{\mathfrak{G}})$, then

13. 10 $$u_n \equiv \dot{E}_\cdot\left[\int_0^{+\infty} e^{-t} f(\dot{x}(\dot{\mathfrak{f}}_n^{-1}))dt\right]$$

tends pointwise under the bound $\|f\|_\infty$ to u. Because $u_n \in D(\dot{\mathfrak{G}}_n)$,

13. 11 $$\mathfrak{Q}u_n = \frac{-e^{u}(db)}{e(db)} = p_n \dot{\mathfrak{G}}_n u_n = p_n(u_n - f)$$

satisfies all the conditions for u_n to belong to $D(\mathfrak{Q})$, and, what is more, $\mathfrak{Q}u_n$ converges pointwise under the bound $2\|f\|_\infty$ to $u - f = \dot{\mathfrak{G}}u$; thus, $u \in D(\bar{\mathfrak{Q}})$ and $\bar{\mathfrak{Q}}u = \dot{G}u$, i.e.,

13. 12 $$\bar{\mathfrak{Q}} \supset \dot{\mathfrak{G}}.$$

As to the proof of $\bar{\mathfrak{Q}} \subset \dot{\mathfrak{G}}$, it is enough to show that if $u \in D(\bar{\mathfrak{Q}})$, then $\dot{G}_1(1 - \bar{\mathfrak{Q}})u = u$, and, for this, it is enough to deduce from $u \in D(\bar{\mathfrak{Q}})$ and $\bar{\mathfrak{Q}}u = u$ that $u \equiv 0$.

Given such a $u \in D(\bar{\mathfrak{Q}})$ with $\bar{\mathfrak{Q}}u = u$ and choosing $u_n \in D(\mathfrak{Q})$ so as to make u_n and $\mathfrak{Q}u_n$ converge pointwise and boundedly to u and $\bar{\mathfrak{Q}}u = u$,

13. 13 $$u_n - h_n = -\int G_D \mathfrak{Q}u de = -\dot{E}_\cdot\left[\int_0^{\dot{\mathfrak{m}}_{\partial D}} (\mathfrak{Q}u_n)(\dot{x})ds\right]$$
$$h_n = \int h_{\partial D}(0, db)u_n(b)$$

implies

13.14 $\dot{E}_\bullet[(u_n-h_n)(\dot{x}(t)),\, t<\dot{\mathfrak{m}}_{\partial D}]-(u_n-h_n)$
$$= \dot{E}_\bullet\left[\int_0^{t\wedge\dot{\mathfrak{m}}_{\partial D}}(\mathfrak{D}u_n)(\dot{x})ds\right] \qquad t\geq 0\,,$$

which, in turn, implies

13.15 $\dot{E}_\bullet[(u-h)(\dot{x}(t)),\, t<\dot{\mathfrak{m}}_{\partial D}]-(u-h)$
$$= \dot{E}_\bullet\left[\int_0^{t\wedge\dot{\mathfrak{m}}_{\partial D}}(\bar{\mathfrak{D}}u)(\dot{x})ds\right] \qquad t\geq 0$$
$$h = \int h_{\partial D}(\cdot,\, db)u(b)\,,$$

and, letting $D\uparrow R^d$ so as to make h tend to a bounded (and hence constant) harmonic function $h(\infty)$,

13.16a $\dot{P}_\bullet[\dot{\mathfrak{m}}_\infty<+\infty] = P_\bullet[\mathfrak{f}(\mathfrak{m}_\infty)<+\infty] = 0 \qquad d=2$

13.16b $u(\infty) = h(\infty) = 0 \qquad d\geq 3$

implies

13.17 $\dot{E}_\bullet[u(\dot{x}(t)]-u = \dot{E}_\bullet\left[\int_0^t u(\dot{x})ds\right] \qquad t\geq 0\,,$

and the desired $u\equiv 0$ follows.

The Green operators leave invariant the space $C(E^d)$ of bounded functions continuous in the ordinary topology of E^d if and only if, for each D, the mean exit time

13.18 $\dot{p} = \dot{E}_\bullet[\dot{\mathfrak{m}}_{\partial D}] = \int G_D de$

is continuous inside D and tends to 0 on ∂D; in this case, the generator $\mathfrak{G}$ coincides with the differential operator $\mathfrak{D}$ acting on the class of functions $u\in C(E^d)$ such that $\mathfrak{D}u\in C(E^d)$; the reader will easily supply the details of the proof.

Here is an example in which the Green operators *do not map $C(E^d)$ into itself.*

Choose $d=3$ and $e=f\times db$, where $f=1+\sum_{n\geq 1}f_n$ and the f_n are the indicators of little non-overlapping open balls D_n converging to as $n\uparrow+\infty$ but not covering 0 itself and so small that

13.19 $P_0[\mathfrak{m}_{\partial D} < -\infty] < 2^{-n} \qquad n \geq 1$.

Because of the first Borel-Cantelli lemma, $p \equiv P.[\mathfrak{m}_{\partial D_n} < +\infty, i.o.] = 0$ at the origin, and it is also clear that $p \equiv 0$ on the rest of R^3. But then $\mathfrak{f} = \int_0^t f(x(s))ds$ is a continuous additive functional, $\mathfrak{f}(s) < \mathfrak{f}(t)$ $(s < t)$, and $x(\mathfrak{f}^{-1})$ is a diffusion with Brownian hitting probabilities and speed measure e.

Given a neighborhood D of 0,

13.20
$$\begin{aligned} & E_0[\min(t : x(\mathfrak{f}^{-1}) \in \partial D)] \\ & = E_0[\mathfrak{f}(\mathfrak{m}_{\partial D})] \\ & \geq \sum_{n\geq 1} E_0\left[\int_0^{\mathfrak{m}_{\partial D}} f_n(x(s))ds\right] \\ & = \sum_{n\geq 1} \int_{D_n} G_D(0, b)db/\text{volume}\,(D_n) \\ & = +\infty\,. \end{aligned}$$

But, as E. B. Dynkin [6] has pointed out, this cannot happen for all small neighborhoods if the Green operators map $C(E^d)$ into itself.

14. DISCONTINUOUS ADDITIVE FUNCTIONALS.

V. A. Volkonskii [16] has studied *discontinuous* additive functionals ; in the present Brownian case their structure is very simple.

A functional $\mathfrak{t}$ *of the Brownian path which satisfies*

14.1 $\mathfrak{t}(t, w)$ is measurable $\boldsymbol{B}_t$ for each $t \geq 0$.

14.2 $0 \leq \mathfrak{t} < +\infty$

14.3 $\mathfrak{t}(t-) = \mathfrak{t}(t)$

14.4 $\mathfrak{t}(t) = \mathfrak{t}(s) + \mathfrak{t}(t-s, w_s^+) \qquad t \geq s$

is the sum of a continuous additive functional $\mathfrak{f}$ *and a discontinuous additive functional* $\mathfrak{j}$ *with*

14.5 $P.[\mathfrak{j}(t) = \mathfrak{j}(0+), t > 0] = 1$

14.6a $P.[\mathfrak{j}(0+) > 0] = 0$ *or* 1

14.6b $C(E) = 0$, *where E is the set of points at which*

$P_\cdot[\mathfrak{j}(0+)>0]=1$ *and* C *is the Newtonian (logarithmic) capacity in* $d\geq 3$ $(=2)$ *dimensions.*[9]

Consider the (discontinuous) additive functional $\mathfrak{j}_n(t)=$ *the sum of the jumps of* $\mathfrak{t}$ *of magnitude* $\geq n^{-1}$ *taking place before time* t, note that $\mathfrak{j}_n(0+)>0)\in \boldsymbol{B}_{0+}$ so that $P_\cdot[\mathfrak{j}_n(0+)>0]=0$ or 1 according to Blumenthal's 01 law, let E_n be the (Borel) set on which $P_\cdot[\mathfrak{j}_n(0+)>0]=1$, and introduce the least *positive* jumping time $\mathfrak{m}$ of $\mathfrak{j}_n$.

If $P_\cdot(\mathfrak{m}<+\infty)>0$ at some point, then

14.7 $$\begin{aligned} 0 &< P_\cdot[0<\mathfrak{m}<+\infty,\ \mathfrak{j}_n(\mathfrak{m})<\mathfrak{j}_n(\mathfrak{m}+)] \\ &= P_\cdot[0<\mathfrak{m}<+\infty,\ x(\mathfrak{m})\in E_n]; \end{aligned}$$

this implies $C(E_n)>0$[10], and it follows that E_n contains a subcompact A of positive capacity, having a (regular) point at which $P_\cdot[x(t)\in A$, i.o., $t\downarrow 0]=1$.[11] But then $P_\cdot[\mathfrak{j}_n(t)\equiv+\infty, t>0]=1$ at that point, contradicting $\mathfrak{j}_n\leq\mathfrak{t}<+\infty$, and it follows that

14.8a $$P_\cdot[\mathfrak{j}_n(t)\equiv\mathfrak{j}_n(0+),\ t>0]\equiv 1$$

14.8b $$C(E_n)=0.$$

The rest is clear: $\mathfrak{j}=\lim_{n\uparrow+\infty}\mathfrak{j}_n$ satisfies 14.5, the remainder $\mathfrak{f}=\mathfrak{t}-\mathfrak{j}$ is a continuous additive functional, 14.6a is immediate from Blumenthal's 01 law, and $C(E)=\lim_{n\uparrow+\infty}C(E_n)=0$.

Massachusetts Institute
of Technology
Kyushu University

---

[9] G. Choquet [3] found that if $E\subset R^d$ is Borel, then

$$\begin{aligned} &\inf C(B) : B \text{ open}, B\supset E \\ &=\sup C(A) : A \text{ compact}, A\subset E; \end{aligned}$$

this common value is the *capacity* of E.

[10] $P_\cdot[x(t)\in E$ *at some positive time*$]$ is positive or $\equiv 0$ according as $C(E)>0$ or not; see, for example, G. Hunt [9].

[11] See O. D. Kellogg [11] for the classical significance of regular points and J. Doob [5] for the probabilistic interpretation.

REFERENCES

[1] R. Blumenthal: An extended Markov property. Trans. Am. Math. Soc. 85 (1957), 52–72.

[2] H. Cartan: Théorie générale du balayage en potentiel newtonien. Ann. Univ. Grenoble 22 (1946), 221–280.

[3] G. Choquet: Theory of capacities. Ann. Inst. Fourier Grenoble, 5 (1953–1956), 131–295.

[4] J. Doob: Stochastic Processes. J. Wiley and Sons. New York (1953).

[5] ———: Semimartingales and subharmonic functions. Trans. Am. Math. Soc. 77 (1954), 86–121.

[6] E. B. Dynkin: Infinitesimal operators of Markov processes. Teor. Veroyat. i Primenen. 1 (1956), 38–60.

[7] ———: Natural topology and excessive functions connected with a Markov process. Doklady Akademii Nayk CCCP, 127 (1959), 17–19.

[8] W. Feller: The general diffusion operator and positivity preserving semi-groups in one dimension. Ann. Math. 60 (1954), 417–436.

[9] G. Hunt: Some theorems concerning Brownian motion. Trans. Am. Math. Soc. 81 (1956), 294–319.

[10] ———: Markov processes and potentials (1). Illinois J. Math. 1 (1957), 44–93.

[11] O. D. Kellogg: Foundations of Potential Theory. J. Springer. Berlin (1929).

[12] K. Itô and H. P. McKean, Jr.: Potentials and the random walk. Illinois J. Math. 4 (1960), 119–132.

[13] ———: Diffusion. to appear.

[14] F. Riesz: Sur les fonctions subharmoniques et leur rapport a la théorie du potentiel. Acta Math. 48 (1926), 329–343.

[15] A. Volkonskiĭ: Random substitution of time in strong Markov process. Teor. Veroyat. i Primenen. 3 (1958), 332–350.

[16] ———: Additive functionals of Markov processes. Doklady Akademii Nayk CCCP. 127 (1959), 735–738.

Note added in proof. The authors came to know that A. D. Ventsel' [19] obtained almost the same result as in sect. 3–5 of the present paper and that A. Meyer [18] studied the same problem for a more general class of Markov processes. As for signed additive functionals, E. B. Dynkin [17] constructed them using stochastic integrals in case of a Brownian motion.

[17] E. B. Dynkin: Additive functionals of a Wiener process determined by stochastic integrals. Teor. Veroyat. i Primenen. 5 (1960), 441–452.

[18] A. Meyer: Representations integrales des potentiels. C. R. Paris 251 (1960), 22 79–2280.

[19] A. D. Ventsel': Non-negative additive functionals of Markov processes. Doklady Akademii Nayk CCCP 137 (1961), 17–20

Z. Wahrscheinlichkeitstheorie 1, 251—257 (1963)

Note on Continuous Additive Functionals of the 1-Dimensional Brownian Path

By

HIROSHI TANAKA

1. Introduction

A functional $\mathfrak{a}(t) = \mathfrak{a}(t, w)$ of $t \geqq 0$ and a Brownian path $w: t \to x(t)$ is called *additive* if

1.1a $\qquad \mathfrak{a}(t, w)$ *depends on* t *and* $x(s): 0 \leqq s \leqq t$ *only,*

1.1b $\qquad \mathfrak{a}(t, w) = \mathfrak{a}(s, w) + \mathfrak{a}(t - s, w_s^+), \quad 0 \leqq s \leqq t,$

where w_s^+ is the *shifted* path $w_s^+: t \to x(t + s)$.

Additive functionals play an important rôle in transforming a Markov process, such as the substitution of time by the inverse functional of a non-negative additive functional (K. ITÔ and H. P. MCKEAN [7], V. A. VOLKONSKII [10]) and the transformation of a Brownian motion to a diffusion with a drift (E. B. DYNKIN [2], M. MOTOO [8]). As for the structure of additive functionals of the several dimensional Brownian motion, A. D. VENTSEL [9] proved that such a functional $\mathfrak{a}$ can be written in the form:

1.2
$$\mathfrak{a}(t, w) = f(x(t)) - f(x(0)) + \int_0^t g(x(s)) \cdot dx(s),$$

under the assumption that $\mathfrak{a}(t, w)$ has a finite expectation for each $t < \infty$. Here, f and g are Borel functions on the state space, g satisfies $\int_0^t g(x(s))^2\, ds < \infty$ $(t < \infty)$ with probability one and the integral in 1.2 is a stochastic integral of K. ITÔ [5].

In this note, it will be proved that, for *every continuous* additive functional $\mathfrak{a}$ of the 1-dimensional Brownian path, $\mathfrak{a}(\tau, w): \tau = \min(t, \mathfrak{m})$ has finite moments of all orders $\geqq 0$, where $\mathfrak{m}$ is the first time $\mathrm{x}(\cdot)$ hits the complement of a bounded interval. This result combined with a slight modification of A. D. VENTSEL's proof implies that *any continuous additive functional of the* 1-*dimensional Brownian path has the representation* 1.2 *with continuous* f.

I wish to thank Professor H. P. MCKEAN for helpful suggestions.

2. Existence of moments

Given the space W of continuous sample paths $w: t \in [0, \infty) \to w_t \in R^1$, write $w_t = x(t, w)$ (or $= x(t) = x_t$ for short), and introducing the corresponding coordinate fields $\boldsymbol{B}_t = \boldsymbol{B}[x(s): s \leqq t]$ and $\boldsymbol{B} = \boldsymbol{B}[x(s): s < \infty]$, let $P_a(\cdot)$ be the *Wiener measure* on $\boldsymbol{B}$ with initial position $a \in R^1$. $[x, \boldsymbol{B}, P.]$ is the so-called (standard) 1-dimensional *Brownian motion*. We use the notation $E_a(f, B)$ for the integral over $B \in \boldsymbol{B}$ of a $\boldsymbol{B}$-measurable function f with respect to P_a and write $E_a(f, B) = E_a(f)$ when $B = W$. The *first passage time* $\inf\{t : x(t) = r\}$ through r and the *first exit time* $\inf\{t : x(t) \notin I\}$ for an interval I are denoted respectively by $\mathfrak{m}_r$ and $\mathfrak{m}_I$. The following *strong Markov property** will be used often: if $\mathfrak{m}$ is a *Markov time***, then for $a \in R^1$ and $B \in \boldsymbol{B}$

$$2.1 \qquad P_a[(P_a(w_{\mathfrak{m}}^+ \in B \mid \boldsymbol{B}_{\mathfrak{m}+}) = P_{x(\mathfrak{m})}(B), \mathfrak{m} < \infty] = P_a(\mathfrak{m} < \infty),$$

where $w_{\mathfrak{m}}^+$ is the *shifted* path $w_{\mathfrak{m}}^+ : t \to x(t + \mathfrak{m})$ and $\boldsymbol{B}_{\mathfrak{m}+}$ is the field generated by the events $B \in \boldsymbol{B}$ such that $B \cap (w : \mathfrak{m} < t) \in \boldsymbol{B}_t$, $t \geqq 0$. By definition, a *continuous additive functional* is a functional $\mathfrak{a}(t, w)$ of t and w which satisfies the following conditions 2.2a—2.2c:

$$2.2\text{a} \qquad \mathfrak{a}(t, w) \textit{ is } \boldsymbol{B}_t\textit{-measurable for each } t \geqq 0,$$

$$2.2\text{b} \qquad |\mathfrak{a}(t, w)| < \infty \text{ and } \mathfrak{a}(t, w) \textit{ is continuous in } t \text{ (a.e.)},$$

$$2.2\text{c} \qquad \mathfrak{a}(t, w) = \mathfrak{a}(s, w) + \mathfrak{a}(t - s, w_s^+), \quad 0 \leqq s \leqq t \text{ (a.e.)},$$

where (a.e.) means (for every w outside a certain set which has P_a-measure zero for all $a \in R^1$).

Now, consider a continuous additive functional $\mathfrak{a}$ of the 1-dimensional Brownian path, and put

$$\bar{\mathfrak{a}}(t, w) = \max_{0 \leqq s \leqq t} |\mathfrak{a}(s, w)|.$$

Then, we have the following

Theorem 1. *For any bounded interval $I = [r_1, r_2]$ we have*

$$2.3 \qquad P_a[\bar{\mathfrak{a}}(\mathfrak{m}) > t] < c_1 e^{-c_2 t}, \quad a \in I,$$

*where $\mathfrak{m} = \mathfrak{m}_I$, and c_1, c_2 are some positive constants depending on r_1 and r_2 but not on $a \in I$. In particular, $\mathfrak{a}(t \wedge \mathfrak{m})$*** has finite moments of all orders $\geqq 0$.*

Proof. First we note that

$$2.4 \qquad -\bar{\mathfrak{a}}(t, w) + \bar{\mathfrak{a}}(s, w_t^+) \leqq \bar{\mathfrak{a}}(t + s, w) \leqq \bar{\mathfrak{a}}(t, w) + \bar{\mathfrak{a}}(s, w_t^+), \quad s, t \geqq 0 \text{ (a.e.)},$$

which follows from the additivity 2.2c of $\mathfrak{a}$. Let $\dot{I}$ be a bounded open interval containing I, and, for each $a \in \dot{I}$, choose a constant $t_a < \infty$ such that $P_a[\bar{\mathfrak{a}}(\dot{\mathfrak{m}}) > t_a] < 1/2$, where $\dot{\mathfrak{m}} = \mathfrak{m}_{\dot{I}}$. Then, using 2.4 and the strong Markov property 2.1, we have

$$\begin{aligned} 1/2 &\geqq P_a[\bar{\mathfrak{a}}(\dot{\mathfrak{m}}) > t_a] \\ &\geqq P_a[\bar{\mathfrak{a}}(\dot{\mathfrak{m}}) > t_a, \mathfrak{m}_b < \dot{\mathfrak{m}}], \qquad b \in \dot{I} \\ &\geqq P_a[\bar{\mathfrak{a}}(\dot{\mathfrak{m}}(w_{\mathfrak{m}_b}^+), w_{\mathfrak{m}_b}^+) - \bar{\mathfrak{a}}(\mathfrak{m}_b, w) > t_a, \mathfrak{m}_b < \dot{\mathfrak{m}}] \\ &\geqq P_a[\bar{\mathfrak{a}}(\dot{\mathfrak{m}}(w_{\mathfrak{m}_b}^+), w_{\mathfrak{m}_b}^+) > 2t_a, \mathfrak{m}_b < \dot{\mathfrak{m}}, \bar{\mathfrak{a}}(\mathfrak{m}_b) < t_a] \\ &= P_a[\mathfrak{m}_b < \dot{\mathfrak{m}}, \bar{\mathfrak{a}}(\mathfrak{m}_b) < t_a] P_b[\bar{\mathfrak{a}}(\dot{\mathfrak{m}}) > 2t_a]. \end{aligned}$$

---

* See G. A. HUNT [4] or R. BLUMENTHAL [1].

** $\mathfrak{m}$ is called a Markov time if $\mathfrak{m} \geqq 0$ and $(w : \mathfrak{m} < t) \in \boldsymbol{B}_t$ for each $t \geqq 0$.

*** $a \wedge b (a \vee b)$ is the smaller (larger) of a and b.

Now, choosing a neighborhood U_a of a such that for all $b \in U_a$

$$P_a[\mathfrak{m}_b < \dot{\mathfrak{m}}, \bar{\mathfrak{a}}(\mathfrak{m}_b) < t_a] > 3/4,$$

it follows that

2.5 $$P_b[\bar{\mathfrak{a}}(\dot{\mathfrak{m}}) > 2t_a] < 2/3, \quad b \in U_a.$$

Because $\bar{\mathfrak{a}}(\mathfrak{m}) \leqq \bar{\mathfrak{a}}(\dot{\mathfrak{m}})$, 2.5 is true for $a \in I$ and $b \in U_a \cap I$ when $\dot{\mathfrak{m}}$ is replaced by $\mathfrak{m}$. Therefore, using Heine-Borel's covering theorem, it is clear that

2.6 $$P_a[\bar{\mathfrak{a}}(\mathfrak{m}) > T] < 2/3, \quad a \in I,$$

for some constant $T < \infty$ independent of a.

Next, put $\sigma_n = \inf\{t : \bar{\mathfrak{a}}(t) \geqq nT\}$ if $\bar{\mathfrak{a}}(t) \geqq nT$ for some $t < \infty$, and $= \infty$ if there is no such t ($n \geqq 1$). Because σ_n ($n \geqq 1$) is a Markov time and

$$\sigma_n \geqq \sigma_{n-1} + \sigma_1(w^+_{\sigma_{n-1}}), \text{ (a.e.)}, \ n \geqq 2,$$

by 2.4, it follows from 2.6 and by induction that

$$\begin{aligned} P_a[\bar{\mathfrak{a}}(\mathfrak{m}) > nT] &= P_a[\sigma_n < \mathfrak{m}] \\ &\leqq P_a[\sigma_{n-1} + \sigma_1(w^+_{\sigma_{n-1}}) < \mathfrak{m}] \\ &= E_a[P_{x(\sigma_{n-1})}(\sigma_1 < \mathfrak{m}), \ \sigma_{n-1} < \mathfrak{m}] \\ &\leqq (2/3)\, P_a(\sigma_{n-1} < \mathfrak{m}) \leqq (2/3)^n, \end{aligned}$$

and this implies 2.3.

3. Representation

Let g be a Borel function on R^1 such that

3.1 $$P_a\Big[\int_0^t g^2(x_s)\,ds < \infty, \ 0 \leqq t < \infty\Big] = 1, \qquad a \in R^1.$$

Then, the stochastic integral $\int_0^t g(x_s)\,dx_s$ is defined (K. Itô [5]).

We first remark that a version of this stochastic integral can be chosen so that it gives a continuous additive functional. When g is bounded, such a version exists, as was discussed by E. B. Dynkin [3]. When g is unbounded, we use the following simple lemma, which corresponds to lemma 1 in [3] and can be proved similarly.

Lemma. *Suppose the stochastic integral* $\int_0^t g_n(x_s)\,dx_s$ *has a version of continuous additive functional* ($n = 1, 2, \ldots$). *Under the notation* $e(a, h) = E_a[\int_0^{\mathfrak{m}_I} h^2(x_s)\,ds]$, *suppose, for each compact interval* I, *that* $e(a, g)$ *and* $e(a, g_n)$ *are finite for* $a \in I$ *and that* $e(a, g - g_n)$ *converges to zero uniformly in* $a \in I$ *as* $n \to \infty$. *Then, the stochastic integral* $\int_0^t g(x_s)\,dx_s$ *has a version of continuous additive functional.*

Given g satisfying 3.1, putting $g_n(b) = g(b)$ for $|g(b)| < n$ and $= 0$ for $|g(b)| \geqq n$, we show that the assumptions of the lemma are satisfied. Theorem 1 applied to the additive functional $\mathfrak{f}(t) = \int_0^t g^2(x_s)\,ds$ implies that $e(a, g)$ is finite, and using the additivity of $\mathfrak{f} \geqq 0$ and the strong Markov property, we see that e is concave and hence continuous on I. By the same reason $e(a, g_n)$ is continuous on I, and since $e(a, g_n)$ increases to $e(a, g)$ as $n \uparrow \infty$ on I, $e(a, g - g_n) = e(a, g) - e(a, g_n)$ tends to zero uniformly on I as $n \to \infty$ Thus the lemma is applicable.

From now on, we take a version of continuous additive functional for a stochastic integral.

Theorem 2. *A continuous additive functional* $\mathfrak{a}$ *of the* 1*-dimensional Brownian path can be written in the following form:*

3.2 $$\mathfrak{a}(t, w) = f(x_t) - f(x_0) + \int_0^t g(x_s)\, dx_s, \quad 0 \leqq t < \infty \qquad \text{(a.e.)},$$

where f *is a continuous function on* R^1 *and* g *satisfies* 3.1.

Because of theorem 1, the method of A. D. VENTSEL is applicable, but here we will give a proof which is different from A. D. VENTSEL's except as regards 3.12. Our method of obtaining the function g in 3.2 seems to be simpler.

Proof. Take $I = [r_1, r_2]$ as before, let $\mathfrak{m} = \mathfrak{m}_I$, and put

3.3a $$\mathfrak{s}(t, w) = \mathfrak{a}(t, w) - [f(x_t) - f(x_0)], \quad t \leqq \mathfrak{m},$$

3.3b $$f = f_I = -E_{\cdot}[\mathfrak{a}(\mathfrak{m})].$$

Denote by $G(a, b)$ the Green function:

$$G(a, b) = 2(a \wedge b - r_1)(r_2 - a \vee b)/(r_2 - r_1),$$

and by $L_2(a)$ $(r_1 < a < r_2)$ the space of those functions on I which are square integrable with respect to the measure $G(a, b)\,db$. $L_2(a)$ is a Hilbert space with inner product $(\Phi_1, \Phi_2)_a = \int_{r_1}^{r_2} G(a, b)\, \Phi_1(b)\, \Phi_2(b)\, db$ for each a; it is independent of $a \in (r_1, r_2)$ as a set. For each $\Phi \in L_2(a)$ let $\mathfrak{s}_\Phi(t \wedge \mathfrak{m})$ be the stochastic integral $\int_0^{t \wedge \mathfrak{m}} \Phi(x_s)\, dx_s$ and note that (see K. ITÔ [5])

3.4a $$E_a[\mathfrak{s}(t \wedge \mathfrak{m})] = E_a[\mathfrak{s}_\Phi(t \wedge \mathfrak{m})] = 0$$

3.4b $$E_a[\mathfrak{s}_{\Phi_1}(\mathfrak{m})\, \mathfrak{s}_{\Phi_2}(\mathfrak{m})] = E_a\Big[\int_0^{\mathfrak{m}} \Phi_1(x_s)\, \Phi_2(x_s)\, ds\Big] = (\Phi_1, \Phi_2)_a.$$

Now, consider the functions $p_\pm(a, \Phi) = E_a[|\mathfrak{s}(\mathfrak{m}) \pm \mathfrak{s}_\Phi(\mathfrak{m})|^2]$ and $p(a, \Phi) = E_a[\mathfrak{s}(\mathfrak{m})\, \mathfrak{s}_\Phi(\mathfrak{m})] = (1/4)(p_+ - p_-)$. Because $\mathfrak{s}$ and $\mathfrak{s}_\Phi$ satisfy 2.2 for $t \leqq \mathfrak{m}$, it follows, using 3.4a and the strong Markov property, that for any compact subinterval $J = [\varrho_1, \varrho_2]$ in (r_1, r_2)

3.5 $$p_\pm(a, \Phi) = E_a[|\mathfrak{s}(\mathfrak{n}) \pm \mathfrak{s}_\Phi(\mathfrak{n})|^2] + E_a[p_+(x(\mathfrak{n}), \Phi)]$$
$$\geqq \frac{\varrho_2 - a}{\varrho_2 - \varrho_1} p_\pm(\varrho_1, \Phi) + \frac{a - \varrho_1}{\varrho_2 - \varrho_1} p_\pm(\varrho_2, \Phi), \quad a \in (\varrho_1, \varrho_2), \quad \mathfrak{n} = \mathfrak{m}_J,$$

and hence $p_\pm$ is concave in (r_1, r_2) and $p_\pm(a, \Phi) \downarrow 0$ as $a \downarrow r_1$ or $a \uparrow r_2$. Therefore, there exists a non-negative measure $\mu_\Phi^\pm$ finite on compact subsets in (r_1, r_2) such that

3.6 $$p_\pm(a, \Phi) = \int_{r_1}^{r_2} G(a, b)\, \mu_\Phi^\pm(db), \quad a \in (r_1, r_2);$$

this $\mu_\Phi^\pm$ is uniquely determined by $p_\pm$; in fact $d\mu_\Phi^\pm = d(-D^+ p_\pm)$*. Hence we have**

3.7 $$p(a,\Phi) = \int_{r_1}^{r_2} G(a,b)\,\mu_\Phi(db), \quad a \in (r_1, r_2)$$
$$d\mu_\Phi = (1/4)\,(d\mu_\Phi^+ - d\mu_\Phi^-) = d(-D^+ p);$$

μ_Φ is the unique signed measure finite on compact subsets in (r_1, r_2) and satisfying 3.7.

Next, we prove that there exists a unique function $g(b)$ independent of a and belonging to $L_2(a)$ such that

3.8 $$p(a,\Phi) = (\Phi, g)_a, \quad \Phi \in L_2(a), \quad a \in (r_1, r_2).$$

This results from the following three steps ((1)—(3)).

(1) If Φ vanishes identically in $J = (\varrho_1, \varrho_2) \subset I$, then μ_Φ has no mass inside J. In fact, in this case $\mathfrak{s}_\Phi = 0$ up to the exist time $\mathfrak{m}_J$ and hence from 3.5

$$p(a,\Phi) = \frac{\varrho_2 - a}{\varrho_2 - \varrho_1}\, p(\varrho_1, \Phi) + \frac{a - \varrho_1}{\varrho_2 - \varrho_1}\, p(\varrho_2, \Phi),$$

i.e., p is linear inside J, proving $d\mu_\Phi = d(-D^+ p) = 0$ there.

(2) If χ_E is the indicator function of a Borel set $E \subset I$, then

3.9 $$p(a, \chi_E) = \int_E G(a,b)\,\mu_{\mathbf{1}}(db),$$

where $\mathbf{1}$ is the function identically equal to one on I. First, from 3.4 b we note that

$$p_+(a,\Phi) + p_-(a,\Phi) = 2p_+(a,\mathbf{0}) + 2\int G(a,b)\,\Phi^2(b)\,db$$

and hence from 3.6 that

3.10 $$d\mu_\Phi^+ + d\mu_\Phi^- = 2\,d\,\mu_{\mathbf{0}}^+ + 2\,\Phi^2\,db.$$

Now, if E is an open interval with $\mu_{\mathbf{0}}^+(\partial E) = 0$ and if $\Phi = \chi_E$, then from (1) and 3.10 $d\mu_\Phi = 0$ on $I - E$ and $d\mu_{\mathbf{1}-\Phi} = 0$ on $\bar{E}$ and hence $d\mu_\Phi = \Phi\, d\mu_{\mathbf{1}}$ follows from the identity $d\mu_\Phi + d\mu_{\mathbf{1}-\Phi} = d\mu_{\mathbf{1}}$. Thus we have 3.9 for such an E. But, because $|p(a,\chi_E)| \leqq \sqrt{E_a[\mathfrak{s}(\mathfrak{m})^2]} \cdot \|\chi_E\|_a$*** by Schwarz's inequality and 3.4 b, $p(a,\chi_E)$ is a signed measure in E. Hence 3.9 must hold for any Borel set E in I because it holds for open intervals E with $\mu_0^+(\partial E) = 0$ and these intervals generate all Borel sets in I.

(3) From (2) we have

3.11 $$p(a,\Phi) = \int G(a,b)\,\Phi(b)\,\mu_{\mathbf{1}}(db), \quad a \in (r_1, r_2),$$

if Φ is a linear combination of finite numbers of indicator functions. On the other hand, for each a, $p(a,\Phi)$ is a linear functional on $L_2(a)$ because of the bound $|p(a,\Phi)| \leqq \sqrt{E_a[\mathfrak{s}(\mathfrak{m})^2]} \cdot \|\Phi\|_a$, and hence by Riesz's theorem there is a unique

---

* D^+ means the right derivative, and $d(-D^+ p_\pm)$ is the measure induced by the function $-D^+ p_\pm$ of bounded variation.

** I owe this reasoning to H. P. McKean.

*** $\|\Phi\|_a = \sqrt{(\Phi, \Phi)_a}$.

$g_a \in L_2(a)$ such that $p(a, \Phi) = (\Phi, g_a)_a$ for $\Phi \in L_2(a)$. Comparing this with 3.11, it follows that $d\mu_1 = g_a db$ and that g_a must be independent of a, which we denote by g. Thus we have 3.8 with a unique $g \in L_2(a)$.

Next, we sketch the proof that $\mathfrak{s}(t) = \mathfrak{s}_g(t)$ for $t \leqq \mathfrak{m}$ (a.e.), following A. D. VENTSEL [9]. Putting $\bar{\mathfrak{s}} = \mathfrak{s} - \mathfrak{s}_g$, it is sufficient to show that

$$E_a[F_1(x(t_1 \wedge \mathfrak{m})) \cdots F_n(x(t_n \wedge \mathfrak{m}))\bar{\mathfrak{s}}(t \wedge \mathfrak{m})] = 0,$$
$$0 \leqq t_1 < \cdots < t_n < \infty, \quad 0 \leqq t < \infty,$$

for any bounded Borel functions $F_1, \ldots, F_n$, and by the additive property of $\bar{\mathfrak{s}}$ and the Markovian property of Brownian motion, it is also enough to prove that

3.12 $$E_a[F(x(t \wedge \mathfrak{m}))\bar{\mathfrak{s}}(t \wedge \mathfrak{m})] = 0$$

for any continuous function F on I. From 3.4b and 3.8 we note that $E_a[\mathfrak{s}_\Phi(\mathfrak{m})\bar{\mathfrak{s}}(\mathfrak{m})] = 0$ and hence by the additive property and 3.4a that

3.13 $$E_a[\mathfrak{s}_\Phi(t \wedge \mathfrak{m})\bar{\mathfrak{s}}(t \wedge \mathfrak{m})] = 0, \quad \Phi \in L_2(a).$$

Now, if F is continuous on I, writing $F = F_0 + h$ where $F_0(r_1) = F_0(r_2) = 0$ and h is a linear function, we have $E_a[h(x(t \wedge \mathfrak{m}))\bar{\mathfrak{s}}(t \wedge \mathfrak{m})] = 0$ by 3.13 and 3.4a. On the other hand, if q is an eigenfunction for the problem $(1/2)q'' = \lambda q$ with $q(r_1) = q(r_2) = 0$, then applying the formula of stochastic integrals (K. ITÔ [6]) to $q(x_t)$ and using 3.13, we have

$$E_a[q(x(t \wedge \mathfrak{m}))\bar{\mathfrak{s}}(t \wedge \mathfrak{m})] = \tfrac{1}{2}\int_0^t E_a[q''(x(s \wedge \mathfrak{m}))\bar{\mathfrak{s}}(s \wedge \mathfrak{m})]\,ds$$
$$= \lambda \int_0^t E_a[q(x(s \wedge \mathfrak{m}))\bar{\mathfrak{s}}(s \wedge \mathfrak{m})]\,ds,$$

and hence $E_a[q(x(t \wedge \mathfrak{m}))\bar{\mathfrak{s}}(t \wedge \mathfrak{m})] = 0$. Now, $E_a[F_0(x(t \wedge \mathfrak{m}))\bar{\mathfrak{s}}(t \wedge \mathfrak{m})] = 0$ follows by approximating F_0 uniformly by a linear combination of q's. Thus 3.12 holds.

Finally, to obtain 3.2, take an increasing sequence of bounded intervals $I_n = [r_{n1}, r_{n2}]$, $n \geqq 1$, with union R^1. We have obtained already

3.14a $$\mathfrak{a}(t, w) = f_n(x_t) - f_n(x_0) + \int_0^t g_n(x_s)\,dx_s, \quad t \leqq \mathfrak{n}_n \qquad \text{(a.e.)},$$

3.14b $$f_n = -E.[\mathfrak{a}(\mathfrak{n}_n)], \quad \mathfrak{n}_n = \mathfrak{m}_{I_n}.$$

Put $l_1(a) = 0$ and for $n \geqq 2$

$$l_n(a) = \sum_{k=1}^{n-1}(r_{k2} - r_{k1})^{-1}[(r_{k2} - a)f_{k+1}(r_{k1}) + (a - r_{k1})f_{k+1}(r_{k2})].$$

Then, because $l_n(x_t) - l_n(x_0)$ can be written as the stochastic integral $\int_0^t l'(x_s)\,dx_s$, 3.14a remains valid when f_n and g_n are replaced respectively by $\bar{f}_n \equiv f_n - l_n$ and $\bar{g}_n \equiv g_n + l_n'$. On the other hand, from 3.14b, it follows that $\bar{f}_n = \bar{f}_{n+1}$ inside I_n and hence $\bar{g}_n = \bar{g}_{n+1}$ inside I_n. Thus, defining $f = \bar{f}_n$ and $g = \bar{g}_n$ inside I_n, we obtain 3.2. 3.1 is clear from the construction of g. Because $\mathfrak{a}$ and the stochastic integral term are continuous in t (a.e.), f must be continuous on R^1. Thus the theorem is completely proved.

Note added in proof. As for the multidimensional case, the author found, in the course of proof reading, A. B. SKOROHOD'S paper (Teor. Verojatn. Primen **6**, 430–439 (1961)) and A. D. VENTSEL'S paper (Doklady Akad. Nauk SSSR n. Ser. **142**, 1223–1226 (1962)). These papers treats the same representation as 3.2 for continuous additive functionals, and especially the latter paper treats the most general continuous addive functionals of a Brownian motion, but the proof given here is different from theirs. The multidimensional version of Theorem 1 was obtained also by H. P. McKEAN (private communication) where I (in the theorem) must be replaced by a suitable *fine* open set.

References

[*1*] BLUMENTHAL, R.: An extended Markoff property. Trans. Amer. math. Soc. **85**, 52—72 (1957).
[*2*] DYNKIN, E. B.: Some transformations of Markov processes. Doklady Akad. Nauk SSSR n. Ser. **133**, 269—272 (1960).
[*3*] — Additive functionals of a Wiener process determined by stochastic integrals. Teor. Verojatn. Primen. **5**, 441—452 (1960).
[*4*] HUNT, G. A.: Some theorems concerning Brownian motion. Trans. Amer. math. Soc. **81**, 294—319 (1956).
[*5*] ITÔ, K.: On stochastic differential equations. Mem. Amer. math. Soc. No. 4.
[*6*] — On a formula concerning stochastic differentials. Nagoya math. J. **3**, 55—65 (1951).
[*7*] —, and H. P. McKEAN: Diffusion. (Forthcoming.)
[*8*] MOTOO, M.: Diffusion process corresponding to $\frac{1}{2} \sum \partial^2/\partial x^{i^2} + \sum b^i(x)\, \partial/\partial x^i$. Ann. Inst. statist. Math. **12**, 37—61 (1960).
[*9*] VENTSEL, A. D.: Additive functionals of several dimensional Wiener process. Doklady Akad. Nauk SSSR n. Ser. **139**, 13—16 (1961).
[*10*] VOLKONSKII, V. A.: Random substitution of time in strong Markov process. Teor. Verojatn. Primen. **3**, 332—350 (1958).

Mathematicae Institute
Kyusyu University
Fukuoka, Japan

(Received February 15, 1962)

Memoirs of the Faculty of Science, Kyushu University
Ser. A, Vol. 18, No. 1, 1964

EXISTENCE OF DIFFUSIONS WITH CONTINUOUS COEFFICIENTS

By
Hiroshi TANAKA

(Received Nov. 30, 1963)

Introduction. The object of this paper is to prove the following theorem. Let a^{ij}, b^i for $1\leq i, j\leq d$ be real-valued continuous functions on Euclidean d-space R^d $(d\geq 2)$, such that for each $x=(x^1, \cdots, x^d)$ in R^d $a^{ij}(x)$ is a positive definite symmetric matrix. Let $B(R^d)$ be the Banach space of all real-valued bounded Borel functions on R^d, with the supremum norm. Then, (1) *there exists a family of linear transformations* T^t *acting on* $B(R^d)$, $0\leq t<+\infty$, *such that*

$$0\cdot 1 \qquad \begin{aligned} &T^tT^s=T^{t+s}, \qquad \|T^t\|\leq 1 \\ &T^tf\geq 0 \quad if \quad f\geq 0 \\ &\lim_{t\downarrow 0} T^tf(x)=f(x), \qquad x\in R^d,\ f\in C(R^d), \end{aligned}$$

and, such that whenever f *is of class* C^2 *with compact support in* R^d,

$$0\cdot 2 \qquad \begin{aligned} &\lim_{t\downarrow 0} t^{-1}\|T^tf(x)-f(x)\| \\ &=\sum_{i,j=1}^{d} a^{ij}(x)\frac{\partial^2 f(x)}{\partial x^i\partial x^j}+\sum_{i=1}^{d} b^i(x)\frac{\partial f(x)}{\partial x^i} \end{aligned}$$

uniformly for x *in* R^d. (2) *There exists a diffusion process on* R^d *with semigroup* $\{T^t\}$.

We also consider the case when b^i $(1\leq i\leq d)$ is bounded measurable. In general, there is no uniqueness for the semigroups satisfying 0.1 and 0.2, and the semigroup constructed here seems to be the smallest one among such semigroups.

If the coefficients a^{ij} and b^i are more smooth, the essentially same problem has been treated in various ways and in somewhat more precise form, such as the (probabilistic) construction of a diffusion based on Ito's stochastic integral equations (b^i and the positive square root of (a^{ij}) are assumed to be

Lipschitz continuous), and the construction of fundamental solutions for parabolic differential equations (the coefficients are assumed to be Hölder continuous). Recently, Ito's stochastic integral equation with coefficients merely continuous or having certain discontinuity was treated by Skorohod [10] and Girsanov [3].

Our construction is based on Hille-Yosida theorem in semigroup theory, and relies heavily on the followings:

(1) A result of Morrey [6] in elliptic differential equations.

(2) McKean's estimate of an a priori kind concerning the continuity of sample paths (4.2 in section 4).

(1) plays an important rôle in the present application of Hille-Yosida theorem. As for the construction of diffusions, we first construct diffusions inside balls, and then obtain a (global) diffusion in R^d connecting these local ones. (2) will be used for proving some properties of local diffusions which will be needed for the later constructions. The present result extends Nelson's result [7] to the case of continuous coefficients.

I wish to thank Professor McKean for helpful talks and for communicating the result (2).

§1. Notations and preliminaries.

All functions we consider are real-valued. We shall use the following notations. If D is a bounded domain in Euclidean d-space $R^d(d\geq 2)$, $\rho(x, \partial D)$ denotes the distance between x and the boundary ∂D of D, and

$$C(D) = \{f:\ f \text{ is bounded continuous in } D\},$$
$$C(\bar{D}) = \{f:\ f \text{ is continuous on the closure } \bar{D} \text{ of } D\},$$
$$C_0(D) = \{f:\ f\in C(D) \text{ and } f(x)\to 0 \text{ as } \rho(x, \partial D)\to 0\},$$
$$C_0(\bar{D}) = \{f:\ f\in C(\bar{D}) \text{ and } f=0 \text{ on } \partial D\}.$$

The supremum norm of a bounded function f on D (or $\bar{D}$) is denoted by $\|f\|_D$ (or $\|f\|_{\bar{D}}$). If n is a positive integer, $C^n(D)=\{f:\ f$ and all its partial derivatives of order $\leq n$ are defined and belong to $C(D)\}$. The first and the second partial derivatives of u is denoted by $u_{,\alpha}$ and $u_{,\alpha\beta}$ respectively ($1\leq\alpha$, $\beta\leq d$). $L_2(D)$ is the real Hilbert space of square summable functions on D. The inner product in $L_2(D)$ is denoted by $(*, *')_{2,D}$ and the norm is by $\|*\|_{2,D}$.

Suppose a second-order elliptic differential operator A:

1·1 $$A=\sum_{i,j=1}^{d} a^{ij}(x)\frac{\partial^2}{\partial x^i \partial x^j}+\sum_{i=1}^{d} b^i(x)\frac{\partial}{\partial x^i},$$

is given in R^d, $x=(x^1, \cdots, x^d)\in R^d$; here we assume that

1·2a a^{ij} is continuous in R^d $(1\leq i, j\leq d)$,

1·2b b^i is measurable and bounded on each compact set in R^d $(1\leq i\leq d)$,

1·2c (a^{ij}) is symmetric positive definite:

$$\sum_{i,j=1}^{d} a^{ij}(x)\lambda_i\lambda_j>0, \qquad (\lambda_1, \cdots, \lambda_d)\neq 0, x\in R^d.$$

Let D be a bounded domain, and denote by $k(s, D)$ the modulus of continuity of (a^{ij}) on D:

1·3 $$k(s, D)=\max_{\substack{|x_1-x_2|\leq s\\ x_1, x_2\in \bar{D}}}\left[\sum_{i,j=1}^{d}|a^{ij}(x_1)-a^{ij}(x_2)|^2\right]^{\frac{1}{2}}, \qquad s>0.$$

We say, for convenience, that

$$A_n=\sum_{i,j=1}^{d} a_n^{ij}(x)\frac{\partial^2}{\partial x^i \partial x^j}+\sum_{i=1}^{d} b_n^i(x)\frac{\partial}{\partial x^i}, \qquad n=1, 2, \cdots,$$

are *approximating operators* of A on D, if the coefficients are defined in some domain $\dot{D}$ containing $\bar{D}$ and satisfy the following conditions 1·4a–d.

1·4a $a_n^{ij}\in C^3(\dot{D})$ and $\|a_n^{ij}-a^{ij}\|_{\bar{D}}\to 0$ as $n\to+\infty$ $(1\leq i, j\leq d)$.

1·4b $b_n^i\in C^3(\dot{D})$, $\|b_n^i\|_{\bar{D}}$ is bounded in $n\geq 1$ and $\|b_n^i-b^i\|_{2, D}\to 0$ as $n\to+\infty$ $(1\leq i\leq d)$.

1·4c If $k_n(s, D)$ is the modulus of continuity of (a_n^{ij}) on D, defined in the same way as 1·3, then

$$k_n(s, D)\leq k(s), \quad s>0, \quad n=1, 2, \cdots,$$

with some function $k(s)$ such that $k(s)\downarrow 0$ $(s\downarrow 0)$.

1·4d (a_n^{ij}) is symmetric and for some constant M

$$0<[\sum_{i,j=1}^{d} a_n^{ij}(x)\lambda_i\lambda_j]^{-1}\leq M, \qquad |\lambda|=1, \quad x\in \bar{D}.$$

Such operators A_n can be obtained as follows. Let φ be a C^∞-function with bounded support such that $\varphi\geq 0$ and $\int_{R^d}\varphi(x)dx=1$ and put

1·5a $\varphi_n(x)=n^d\varphi(nx)$,

1·5b $a_n^{ij}(x)=(a^{ij}*\varphi_n)(x)=\int a^{ij}(x-y)\varphi_n(y)dy$, $b_n^i(x)=(b^i*\varphi_n)(x)$.

Then, if the support of φ is contained in a sphere of sufficiently small radius with center the origin, the a_n^{ij} and b_n^i have all the properties required for the coefficients of approximating operators. The function $k(s)$ in 1·4c can be taken as $k(s, \dot{D})$.

The following is a result of Morrey [6: Theorem 4.7], and is quoted here in a simple and restricted form for the convenience of the present applications.

Let D be a bounded domain, A an elliptic differential operator in D with the form 1·1, and assume that

1·6a $a^{ij}\in C(\bar{D})$ and b^i is bounded measurable $(1\leq i, j\leq d)$,

1·6b there exist constants M, M_1 and M_2 such that

$$\|a^{ij}\|_{\bar{D}}\leq M_1, \qquad \|b^i\|_{\bar{D}}\leq M_2, \qquad (1\leq i, j\leq d),$$
$$0<[\sum_{i,j}^{d} a^{ij}(x)\lambda_i\lambda_j]\leq M, \qquad |\lambda|=1, x\in\bar{D}.$$

Then, 1) *if u has continuous partial derivatives up to the second order, and if u and Au belong to $L_2(D)$, then*

1·7 $$[\sum_{\alpha,\beta=1}^{d}\|u_{,\alpha\beta}\|_{2, D_0}^2]^{\frac{1}{2}}\leq L_1, \qquad \bar{D}_0\subset D,$$

where the constant L_1 depends only on

1·8 *the dimension d, M, M_1, M_2, D_0, $\|u\|_{2,D}$, $\|Au\|_{2.D}$, the modulus of continuity $k(s, D)$ of (a^{ij}) on D, and the distance of D_0 from ∂D.*

2) *If, in addition to the hopotheses in* 1), *Au satisfies*

1·9 $$\|Au\|_{2.B(x,r)}\leq L(r/\rho)^{\lambda}, \qquad 0\leq r\leq\rho=\rho(x, \partial D),$$

with $\lambda=(d/2)-1+\mu, 0<\mu<1$, for all $B(x, r)\equiv\{y:|y-x|<r\}\subset D, x\in D$, then

1·10 $$\|u\|_{\bar{D}_0}\leq L_2, \quad \|u_{,\alpha}\|_{\bar{D}_0}\leq L_3 \qquad (1\leq\alpha\leq d),$$
$$|u_{,\alpha}(x)-u_{,\alpha}(y)|\leq L_4|x-y|^{\mu}, \qquad x, y\in\bar{D}_0\subset D,$$

where the constants L_2, L_3 and L_4 depend only on the quantities 1·8 *and on*

μ and L.

In particular, if Au is bounded, 1·9 is satisfied with $L=\omega^{1/2}\rho_0^{d/2}\|Au\|_D$ and $\mu=1/2$ where ω is the volume of the unit ball in R^d and $\rho_0=$(diameter of D)/2.

§ 2. Extension of A.

Given A with the coefficients satisfying 1·2, we shall define an extension $\bar{A}$ of A. First, for each $x_0\in R^d$, we denote by $\mathfrak{D}_{x_0}(\bar{A})$ the collection of all functions u defined in some neighborhoods (which may depend on each u) of x_0 with the following property: for each $u\in\mathfrak{D}_{x_0}(\bar{A})$, there exist a bounded domain U containing x_0, approximating operators A_n of A on U and a sequence of functions $u_n\in C^2(U)$ such that u_n and $A_nu_n\equiv f_n$ converges uniformly in U to u and a certain limit $f\in C(U)$ respectively. We want to define $\bar{A}$ by $\bar{A}u(x_0)=f(x_0)$ for $u\in\mathfrak{D}_{x_0}(\bar{A})$. But, for this, we must prove that the above limit f does not depend on the choice of (U, A_n, u_n).

To prove this, suppose there exist two (U, A_n, u_n) and (U', A_n', u_n') with the property in the definition of $\mathfrak{D}_{x_0}(\bar{A})$. In U we consider the equation

2·1 $$A_nu_n=f_n\,.$$

For this equation, as is easily verified, all the conditions assumed in Morrey's theorem are satisfied; in fact, the corresponding quantities 1·8, L and μ can be chosen to be independent of n, because the coefficients of A_n satisfy 1·4 and $\|u_n\|_D$ and $\|f_n\|_D$ are bounded in $n\geq1$. Thus, if D_0 is a domain such that $D_0\ni x_0$ and $\bar{D}_0\subset U\cap U'$, Morrey's theorem applied to the equation 2·1 implies the existence of constants L_1, L_2, L_3 and L_4 independent of n such that 1·7 and 1·10 hold for $u=u_n$, $n\geq1$. Hence, we have

2·2 $$\sup_{n\geq1}[\|u_n\|^2_{2,D_0}+\sum_{\alpha=1}^{d}\|u_{n,\alpha}\|^2_{2,D_0}+\sum_{\alpha,\beta=1}^{d}\|u_{n,\alpha\beta}\|^2_{2,D_0}]^{\frac{1}{2}}<+\infty\,.$$

By the same reason, 2·2 remains true when all the u_n are replaced by u_n'. Now, introduce in $C^2(D_0)$ the inner product:

$$\langle u, v\rangle=(u, v)_{2,D_0}+\sum_{\alpha}(u_{,\alpha}, v_{,\alpha})_{2,D_0}+\sum_{\alpha,\beta}(u_{,\alpha\beta}, v_{,\alpha\beta})_{2,D_0}\,,$$

and denote by $\mathbf{H}$ the real Hilbert space obtained by the completion of $C^2(D_0)$ with respect to the metric $|||u-v|||\equiv\langle u-v, u-v\rangle^{\frac{1}{2}}$. Then, 2·2 means that

$\{u_n, n\geq 1\}$ is a bounded set in H, and therefore contains a subsequence converging weakly. By Banach-Saks theorem, one can choose a sequence of certain convex combinations v_k of elements in that subsequence so that v_k converges strongly in **H**. v_k can be written as follows:

$$v_k=\sum_{l=k}^{l_k} c_{kl}u_l, \qquad c_{kl}\geq 0, \qquad \sum_{l=1}^{l_k} c_{kl}=1\,.$$

By the same argument, there exist some convex combinations v'_k:

$$v'_k=\sum_{l=k}^{l'_k} c'_{kl}u'_l\,, \qquad c'_{kl}\geq 0, \qquad \sum_{l=k}^{l'_k} c'_{kl}=1\,,$$

converging strongly. Put $\tilde{v}_k=v_k-v'_k$. Then, each of $\tilde{v}_k$, $\tilde{v}_{k,\alpha}$ and $\tilde{v}_{k,\alpha\beta}$ $(1\leq\alpha, \beta\leq d)$ converges in the norm of $L_2(D_0)$. But, because by the assumption u_n and u'_n have the same uniform limit u on D_0, the $L_2(D_0)$-limit of $\tilde{v}_k$ is zero, and hence the same holds for $\tilde{v}_{k,\alpha}$ and $\tilde{v}_{k,\alpha\beta}$ as is verified by taking Fourier transform. Hence one has $\|A\tilde{v}_k\|_{2,D_0}\to 0$ as $k\to+\infty$. On the other hand, using the assumptions on u_n and u'_n and on the approximating operators it follows that

$$A\tilde{v}_k=\sum_{l=k}^{l_k} c_{kl}A_l u_l-\sum_{l=k}^{l'_k} c'_{kl}A'_l\,u'_l+\sum_{l=k}^{l_k} c_{kl}(A-A_l)u_l-\sum_{l=k}^{l'_k} c'_{kl}(A-A'_l\,)u'_l$$

converges to $f-f'$ $(f'=\lim A'_n u'_n)$ in the norm of $L_2(D_0)$. Thus $f=f'$ almost everywhere on D_0, and hence $f(x)=f'(x)$ for $x\in D_0$ by the continuity.

We now define $\bar{A}$ by $\bar{A}u(x)=f(x)$, $x\in U$ for $u\in\mathfrak{D}_{x_0}(\bar{A})$ where U and f are the same ones in the definition of $\mathfrak{D}_{x_0}(\bar{A})$. $\mathfrak{D}_{x_0}(\bar{A})$ is the local domain at x_0 of $\bar{A}$. Note that if u is defined on a domain D and $u\in\mathfrak{D}_x(\bar{A})$ for every $x\in D$, then $\bar{A}u(x)$ is continuous in D.

LEMMA 2·1. *If* $u\in\mathfrak{D}_{x_0}(\bar{A})$ *and has a local maximum at* x_0 *then* $\bar{A}u(x_0)\leq 0$.

Proof. The proof is similar to that of Lemma 1 of Ventsel [13]. Suppose, on the contrary, that $\bar{A}u(x_0)>0$. By the definition of $\mathfrak{D}_{x_0}(\bar{A})$, there exist a bounded domain U containing x_0, $u_n\in C^2(U)$ $(n\geq 1)$ and approximating operators A_n such that u_n and $A_n u_n$ converge to u and $\bar{A}u$ uniformly on U respectively. Then, there exist ε, $\delta>0$ such that $A_n u_n(x)>\varepsilon$ for $|x-x_0|\leq\delta$ for all sufficiently large n. Put

$$\bar{u}_n(x)=u_n(x)-\varepsilon[\max_{|x-x_0|\leq\delta} A_n|x-x_0|^2]^{-1}|x-x_0|^2.$$

Then for all sufficiently large n, $A_n\bar{u}_n(x)>0$ for $|x-x_0|\le\delta$, and hence, $\bar{u}_n$ can not have a local maximum inside $|x-x_0|<\delta$. Thus

$$u_n(x_0)=\bar{u}_n(x_0)<\max_{|x-x_0|=\delta}\bar{u}_n(x)$$
$$=\max_{|x-x_0|=\delta}u_n(x)-\varepsilon\delta^2[\max_{|x-x_0|\le\delta}A_n|x-x_0|^2]^{-1},$$

and hence

$$u(x_0)\le\max_{|x-x_0|\le\delta}u(x)-2\varepsilon\delta^2[\sum_{i=1}^{d}\{\|a_n^{ij}\|_{B(x_0,\delta)}+\delta\|b_n^i\|_{B(x_0,\delta)}\}]^{-1}.$$

This is true for all sufficiently small $\delta>0$, and therefore contradicts the hypothesis, proving $\tilde{A}u(x_0)\le 0$.

§ 3. The equation $(\lambda-\bar{A})u=f$ and the semigroup on $C_0(\bar{D})$.

In this and later sections, D denotes an open ball $B(a, r)$ with center a and radius r. Take approximating operators $\{A_n\}$ of A on D, and let us prepare a simple lemma.

LEMMA 3·1. *Let $0<r_0\le r$ and $D_0=\{x:|x-a|<r_0\}$. Let u_n, $n\ge 1$, be the solution of*

3·1
$$\begin{cases} A_nu_n=-f & \text{in } D_0,\ f\in C(\bar{D}_0)\\ u_n=0 & \text{on } \partial D_0\,. \end{cases}$$

Then, there exists a function $\omega(s)$ $(s>0)$, independent of $n\ge 1$ and r_0, such that $\omega(s)\downarrow 0$ as $s\downarrow 0$ and $|u_n(x)|\le\|f_n\|_{\bar{D}_0}\cdot\omega(\rho)$, $\rho=\rho(x, \partial D)$.

Proof. It is sufficient to prove the lemma in the case $f\equiv 1$, because $|u_n(x)|\le\|f\|_{\bar{D}_0}\cdot|u_n'(x)|$ where u_n' is the solution of 3·1 corresponding to $f_n=1$. Take any $x_0\in\partial D_0$, and let $l=(l_{pq})$ be an orthogonal transformation which sends the vector $r_0^{-1}(a-x_0)$ to the unit vector $(1, 0, \cdots, 0)$. Let M be the constant in 1·4d and $M_2=\sup\|b_n^i\|_{\bar{D}}$, and consider the function

$$v_{x_0}(x)=c_1^2c_2e^{2c_1r_0}\{1-e^{-2c_1l_1(x-x_0)}\}-c_1^{-1}c_2l_1(x-x_0),$$

$$l_1(x-x_0)=\sum_{p=1}^{d}l_{1p}(x^p-x_0^p)\,,\qquad c_1=\sqrt{d}\cdot M\cdot M_2,\qquad c_2=2M\,.$$

Then, using the orthogonality of l and 1·4d, it follows that

3·2
$$v_{x_0}(x)\ge 0\quad\text{and}\quad A_nv_{x_0}(x)\le -2\,,$$

for x in $\{y: 0\leq l_1(y-x_0)\leq 2r_0\}$ and hence in D_0. Thus

3·3 $$A_n(v_{x_0}-u_n)\leq -1$$

in D_0. We now prove that $u_n\leq v_{x_0}$ on $\bar{D}$. Suppose it were not true, and let x_1 be a point at which $v_{x_0}-u_n$ attains its negative minimum on $\bar{D}_0$. Then x_1 must be inside D_0, because $v_{x_0}-u_n\geq 0$ on ∂D_0 by 3·1 and 3·2. Therefore, $A_n(v_{x_0}-u_n)(x_1)\geq 0$ and this contradicts 3·3. Since $u_n\geq 0$, one has $0\leq u_n\leq v_{x_0}$. Thus the lemma holds with

$$\omega(s)=c_1^{-2}c_2e^{2c_1r}(1-e^{-2c_1s})+c_1^{-1}c_2s\,.$$

THEOREM 3·1. *For any $\lambda>0$ and $f\in C(\bar{D})$ the equation*

3·4 $$(\lambda-\bar{A})u=f \quad \text{in } D, \quad u=0 \quad on \quad \partial D\,,$$

is uniquely solvable in $C_0(\bar{D})$, and the solution u satisfies

3·5 $$\|u\|_{\bar{D}}\leq\lambda^{-1}\|f\|_{\bar{D}}, \quad and \quad u\geq 0 \quad if \quad f\geq 0\,,$$

and is the uniform limit of the sequence of the solutions u_n of

3·6 $$(\lambda-\bar{A}_n)u=f \quad in\ D, \quad u=0 \quad on \quad \partial D.$$ [1]

Proof. We first notice that any solution of 3·4 or 3·6 satisfies 3·5, which, in fact, follows from Lemma 2·1. If f is of class C^1 on $\bar{D}$, then it is well-known that the solution $\in C^2(D)$ of $(\lambda-A_n)u=f$ in D and $u=0$ on ∂D exists uniquely (for example, see [5: p. 84]). For general $f\in C(\bar{D})$, it follows from the bound 3·5 that the solution u_n of 3·6 is the uniform limit (on $\bar{D}$ and in $n\geq 1$), as $k\uparrow +\infty$, of the solutions $u_{n,k}$ of $(\lambda-A_n)u=f_k$ in D and $u=0$ on ∂D, where f_k is a sequence of functions of class C^1 on $\bar{D}$ converging to f uniformly on $\bar{D}$. We now prove that $u_{n,n}$ converges to the solution 3·4 uniformly as $n\uparrow +\infty$. Because of the bound 3·5 for $u_{n,n}$, 1·9 with Au replaced by $\lambda u_{n,n}-f_n$ holds with constants L and μ independent of n, while A_n satisfies 1·4. Therefore, the Morrey's theorem applied to the equation $A_nu_{n,n}=\lambda u_{n,n}-f_n$ implies that $u_{n,n}$ and its all first derivatives are bounded, on each subdomain D_0 with $\bar{D}_0\subset D$, by certain constants independent of n; in particular, $\{u_{n,n}\}$ constitutes a family of equicontinuous functions on $\bar{D}_0$. Using Lemma 3·1 and noting that D_0 is arbitrary under the condition $\bar{D}_0\subset D$, $\{u_{n,n}\}$ turns out to be an equicontinuous family on $\bar{D}$. Thus, there exists a subsequence of

1) $\bar{A}_n$ is related to A_n as $\bar{A}$ is to A. Similar notations with suffix n will be used throughout.

$\{u_{n,n}\}$ which converges to a certain limit u uniformly on $\bar{D}$. This limit u solves 3·5 because $u\in C_0(\bar{D})\cap\{\bigcap_{x\in D}\mathfrak{D}_x(\tilde{A})\}$, and the uniqueness of the solution follows from Lemma 2·1. Finally, because $u_{n,k}$ converges to u_n uniformly on $\bar{D}$ and in $n\geq 1$ as $k\uparrow+\infty$, and because every (uniform) limit function in the (compact) set $\{u_{n,n}\}$ is the unique solution of 3·4, u_n converges to u uniformly on $\bar{D}$ as $n\uparrow+\infty$. Q.E.D.

Denote by $\mathfrak{D}(\tilde{A}_D)$ the space of functions $u\in C_0(D)$ such that $u\in\mathfrak{D}_x(\tilde{A})$ for each $x\in D$ and $\tilde{A}u\in C_0(D)$. We define $\tilde{A}_D$ by $\tilde{A}_Du(x)=\tilde{A}u(x)$, $x\in D$, for $u\in\mathfrak{D}(\tilde{A}_D)$.

THEOREM 3·2. *If $\mathfrak{D}(\tilde{A}_D)$ is dense in $C_0(D)$, then* (1) *there exists a unique semigroup of operators $\{T_D^t, t>0\}$ acting on $C_0(D)$ with generator $\tilde{A}_D$ such that*

3·7 $$\lim_{t\downarrow 0}\|T_D^tf-f\|_D=0,\qquad f\in C_0(D)$$

3·8 $$T_D^tf\geq 0\quad \textit{if}\quad f\geq 0,\quad \textit{and}\quad \|T_D^t1\|_D\leq 1\,.$$

(2) *Let $\{T_{n,D}^t\}$ denote the semigroup with generator $\tilde{A}_{n,D}$. Then*

3·9 $$\lim_{n\uparrow+\infty}\|T_D^tf-T_{n,D}^tf\|_D=0,\qquad f\in C_0(D),\quad t>0\,.$$

Proof. Under the assumption, Theorem 3·1 means that, for each $\lambda>0$, $(\lambda-\tilde{A}_D)$ maps the dense subset $\mathfrak{D}(\tilde{A}_D)$ of $C_0(D)$ onto $C_0(D)$ in one to one way, and the inverse $(\lambda-\tilde{A}_D)^{-1}$ is a positive operator with norm $\leq\lambda^{-1}$. Thus, (1) follows from Hille-Yosida's theorem. (2) results from Theorem 3·1 using Trotter [12: Theorem 5·1]. Q.E.D.

Remark. If b^i is continuous $(1\leq i\leq d)$, then $\mathfrak{D}(\tilde{A}_D)$ is dense in $C_0(D)$, because $\mathfrak{D}(\tilde{A}_D)$ contains functions in $C^2(D)$ with compact supports $\subset D$. Without the continuity of b^i, this dense property of $\mathfrak{D}(\tilde{A}_D)$ will be proved in the next section.

§4. $\{T_{n,D}^t\}$ and the corresponding diffusion.

To the semigroup $\{T_{n,D}^t\}$ there corresponds a unique system of transition measures $\{P_{n,D}(t, x, \cdot),\ t>0,\ x\in D\}$ on D with total mass ≤ 1 such that

$$T_{n,D}^tf(x)=\int f(y)P_{n,D}(t, x, dy),\qquad f\in C_0(D)\,.$$

These transition measures give rise to a diffusion on $\tilde{D}$. To be precise, let $\tilde{D}$

be the one-point compactification $D\cup\{\infty\}$ of D (∞ is an extra point), and introduce the sample space $\mathbf{W}_D$ of continuous paths w: $t\in[0,+\infty)\to x(t,w)$ $=x(t)=x_t\in D$ such that $x(t,w)\in D$ for $t<\sigma_\infty$ and $=\infty$ for $t\geq\sigma_\infty$ where $0\leq\sigma_\infty(w)\leq+\infty$, and the corresponding coordinate field $\mathbf{B}_D$. Then, probability measures $\mathbf{P}^x_{n,D}(\cdot)$, $x\in D$, on $\mathbf{B}_D$ can be introduced by

$$\mathbf{P}^x_{n,D}[x(t)\in A]=\mathbf{P}_{n,D}(t,x,A),\quad x\in D,\ A\subset D,$$
$$\mathbf{P}^\infty_{n,D}[x(t)=\infty,\quad t\geq 0]=1.$$

$[\mathbf{W}_D, \mathbf{B}_D, \mathbf{P}^x_{n,D}, x\in D]$ is a strong Markov process[2], which is called $\bar{A}_{n,D}$-diffusion.

$\bar{A}_{n,D}$-diffusion can also be constructed using the stochastic integral equation of K. Ito [4] as follows. Let $\alpha^{ij}_n(x)$ be the positive definite square root of the matrix $\{2a^{ij}_n(x)\}$, extend $\alpha^{ij}_n(x)$ and $b^i_n(x)$ $(1\leq i,j\leq d)$ to the whole of R^d so that they satisfy the Lipschitz condition, and let $y^{(x)}_t$ be the solution of the stochastic integral equation:

$$4\cdot 1 \qquad y^{(x)i}_t=x^i+\sum_{j=1}^d\int_0^t\alpha^{ij}_n(y^{(x)}_s)d\xi^j_s+\int_0^t b^i_n(y^{(x)}_s)ds\,,$$

where $\xi_t=(\xi^1_t,\cdots,\xi^d_t)$ is a Brownian motion in R^d. Then

$$\mathbf{P}^x_{n,D}(x_t\in A)=\mathbf{P}(y^{(x)}_t\in A,\quad \tau<t),\qquad x\in D,\quad A\subset D;\quad \tau=\inf\{t\geq 0: y^{(x)}_t\notin D\}.$$

Using the above stochastic integral representation of the sample path, H. P. McKean[3] showed that for any open set U and compact set $F\subset U$

$$4\cdot 2\text{a} \qquad \mathbf{P}^x_{n,D}(x_t\in U^c)=o(t),\qquad t\downarrow 0\,,$$

$o(t)$ being uniform in $x\in F$ and in $n\geq 1$, and that for any $\varepsilon>0$ there exists a constant $\delta>0$ independent of n such that

---

2) Strong Markov property results from the fact that the corresponding semigroup maps $C_0(D)$ into itself.

3) 4.2 are the consequences of the following McKean's sharper estimate concerning stochastic integrals: if $(\alpha^{ij}(t))$ is a positive definite $d\times d$-matrix with eigenvalues bounded above and below by positive constants, and each component is a measurable function of the sample path $\{\xi^i(s); s\leq t\}$ of the d-dimensional Brownian motion (ξ^i), and if $\boldsymbol{\alpha}(t)$ is the d-dimensional vector with components $\sum_{j=1}^d\int_0^t\alpha^{ij}(s)\xi^j(ds)$ (stochastic integrals), then

$$P[|\boldsymbol{\alpha}(t)|>\varepsilon]\leq K\cdot e^{-\frac{c}{t}}$$

where K, $c>0$ depend only on the bounds of eigenvalues, $\varepsilon>0$ and d.

4·2b $$\mathbf{P}^x_{n,D}[\sup_{\substack{0<t_1,t_2<\sigma_\infty \\ |t_2-t_1|\leq\delta}} |x(t_2)-x(t_1)|<\varepsilon]>1-\varepsilon\,.$$

Let $G^\lambda_{n,D}$ be the Green operator:

4·3 $$G^\lambda_{n,D}f(x)=\mathrm{E}^x_{n,D}\left[\int_0^{\sigma_\infty} e^{-\lambda t}f(x_t)dt\right], \qquad \lambda>0^{4)},$$
$$=\int_0^{+\infty} e^{-\lambda t}T^t_{n,D}f(x)dt, \qquad \text{if} \quad f\in C_0(D)\,.$$

Then, $u=G^\lambda_{n,D}f$ ($f\in C_0(D)$) is the solution of 3·6. 4·1 implies that

4·4 $$\|\lambda G^\lambda_{n,D}f-f\|_D\to 0, \qquad f\in C_0(D)\,,$$

uniformly in $n\geq 1$ as $\lambda\uparrow+\infty$. But, because by Theorem 3·1 for each $\lambda>0$ $G^\lambda_{n,D}f$ converges uniformly on D to the solution u_λ of 3·4, 4·4 holds when $G^\lambda_{n,D}f$ is replaced by u_λ, and in particular, *$\mathfrak{D}(\bar{A}_D)$ is dense in $C_0(D)$*. Thus, Theorem 3·2 *is valid without the continuity assumption of b^i*.

§ 5. Construction of the diffusion with generator $\bar{A}$.

Let $\{T^t_D\}$ be the semigroup constructed in Theorem 3·2. By the Riesz-Markoff theorem, for each $x\in D$ and $t>0$ there is a unique regular positive measure $P_D(t, x, \cdot)$ with total mass ≤ 1 such that

$$T^t_D f(x)=\int f(y)P_D(t, x, dy), \qquad f\in C_0(D)\,,$$
$$P_D(t+s, x, \cdot)=\int P_D(s, x, dy)P_D(t, y, \cdot)\,.$$

Defining

$$\begin{aligned} &P_D(t, x, \infty)=1-P_D(t, x, D), && x\in D \\ &P_D(t, \infty, A)=1 &&, \quad \text{if} \quad A\ni\infty \\ &\qquad\qquad\quad\;\; =0 &&, \quad \text{otherwise,} \end{aligned}$$

$\{P_D(t, x, \cdot), x\in\tilde{D}\}$ turns out to be a transition probability on $\tilde{D}$.

LEMMA 5·1. *$P_D(t, x, \cdot)$ has a uniform local character: if U is open in $\tilde{D}$ and F is a closed subset of U then $t^{-1}P_D(t, x, U^c)\to 0$ uniformly in $x\in F$ as $t\downarrow 0$.*

---

4) $\mathrm{E}^x_{n,D}[f\,;B]=\int_B f(w)\mathrm{P}^x_{n,D}(dw)$ and $\mathrm{E}^x_{n,D}[f]=\mathrm{E}^x_{n,D}[f\,;\mathrm{W}_D]$. We also use the similar notations without suffix n.

Proof. If $U \not\ni \infty$, take another open set U_1 such that $F \subset U_1$ and $\bar{U}_1 \subset U$, and choose $f \geq 0$ in $C_0(D)$ such that $f=1$ in U_1 and $=0$ outside U. Then, from 3·9 and 4·2a,

$$\begin{aligned} P_D(t, x, U^c) &\leq 1 - T_D^t f = 1 - \lim_{n \to +\infty} T_{n,D}^t f \\ &\leq 1 - \varliminf_{n \to +\infty} P_{n,D}(t, x, U_1) \\ &\leq \varlimsup_{n \to +\infty} P_{n,D}^x[x(t) \in U_1^c] = o(t), \end{aligned}$$

$o(t)$ being uniform in $x \in F$. The case when $U \ni \infty$ is treated similarly.

By this lemma, using Ray's theorem[5], one can introduce a family of probability measures $\{\mathbf{P}_D^x(\cdot),\ x \in \tilde{D}\}$ on $\mathbf{B}_D$ so that $[\mathrm{W}_D, \mathbf{B}_D, \mathbf{P}_D^x, x \in \tilde{D}]$ is a diffusion process[5] with transition probability $P_D(t, x, \cdot)$. We call this process $\bar{A}_D$-diffusion. Of course, $G_D^\lambda f$ ($f \in C_0(D)$) defined similarly to 4·3 is the solution of 3·4.

Next, take an open ball D_1 such that $D_1 \supset \bar{D}$. Then, if $\{A_n\}$ are approximating operators on D_1, they are also approximating operators on D and $\bar{A}_{n,D}$-diffusion up to σ_∞ is identical in law to $\bar{A}_{n,D_1}$-diffusion up to the exit time $\sigma_D = \inf\{t : x(t) \notin D\}$, because for $f \in C_0(D)$

$$u_n(x) = \mathbf{E}_{n,D_1}^x\left[\int_0^{\sigma_D} e^{-\lambda t} f(x_t) dt\right], \qquad \lambda > 0, \quad n \geq 1 \tag{5·1}$$

solves the (resolvent) equation 3·6 for $\bar{A}_{n,D}$-diffusion. But, this identity in law is not obvius for $\bar{A}_D$- and $\bar{A}_{D_1}$-diffusions. Let $\tilde{W}_{D_1}$ be the space of continuous paths

$$w : [0, +\infty) \to x(t, w) = x(t) = x_t \in \hat{D}_1 = D_1 \cup \{\infty\}$$

such that $\lim_{t \to +\infty} x(t, w) = \infty$. $\tilde{W}_{D_1}$ endowed with the metric

$$\rho(w, w') = \sup_{0 \leq t < +\infty} \tilde{\rho}(x(t, w), x(t, w'))$$

is a complete metric space, where $\tilde{\rho}$ is the metric of the compact set $\hat{D}_1$. Each of $\mathbf{P}_{n,D_1}^x(\cdot)$ and $\mathbf{P}_{D_1}^x(\cdot)$ is a probability measure on the topological Borel field of $\tilde{W}_{D_1}$, and because of 3·9 and 4·2b an application of Prohorov's theorem [8:p. 181] yields, for any bounded continuous function F on $\tilde{W}_{D_1}$,

---

5) See [9] or [2:Chap. 6].

5·2 $$\lim \mathrm{E}^x_{n,D_1}(F)=\mathrm{E}^x_{D_1}(F)\,.$$

LEMMA 5·2. *For any* $\varepsilon>0$ *and* $x\in D$ *there exist continuous functions* h_1 *and* h_2 *on* $\tilde{W}_{D_1}$ *such that* $0\le h_1\le\sigma_D\le h_2$ *and* $\mathrm{E}^x_{n,D_1}(h_2-h_1)<\varepsilon$.

Proof. Take $D_2=B(a, r_2)$ and $D_0=B(a, r_0)$ so that $x\in D_2$, $\bar{D}_2\subset D$, $\bar{D}\subset D_0$, $\bar{D}_0\subset D_1$; put $\sigma_0=\sigma_{D_0}$, $\sigma_2=\sigma_{D_2}$. Because $\mathrm{E}^x_{n,D_1}(\sigma_0)$ is equal to

$$\mathrm{E}^x_{n,D_1}(\sigma_2)+\mathrm{E}^x_{n,D_1}[\mathrm{E}^{x(\sigma_2)}_{n,D}(\sigma_0)]$$

and is the solution of 3·1 with $f=1$, using Lemma 3·1 one sees that there exist r_0 and r_2 such that $\mathrm{E}^x_{n,D}(\sigma_0-\sigma_2)<\varepsilon$, $n\le 1$. Let g_1 and g_2 be continuous functions on $[0,+\infty)$ such that $0\le g_1$, $g_2\le 1$, $g_1=1$ on $[0, r_2]$ and $g_1=0$ on $[r,+\infty)$, $g_2=1$ on $[0, r]$ and $g_2=0$ on $[r_0,+\infty)$. Then the following functions have the desired properties.

$$h_i(w)=\int_0^{+\infty} g_i(\max_{0\le s\le t}|x_s-a|)ds, \qquad i=1, 2.$$

LEMMA 5·3. *$\tilde{A}_D$-diffusion up to σ_∞ is identical in law to $\tilde{A}_{D_1}$-diffusion up to σ_D.*

Proof. It is enough to prove that for $f\ge 0$ in $C_0(D_1)$

$$u(x)=\mathrm{E}^x_{D_1}\left[\int_0^{\sigma_D} e^{-\lambda t}f(x_t)dt\right], \qquad \lambda>0,\quad x\in D,$$

is the limit of $u_n(x)$ defined by 5.1 as $n\uparrow+\infty$, because u_n is the solution of 3·6 and hence by Theorem 3·1 converges to the solution $G^\lambda_D f$ of 3·4. Fix $x\in D$, and for given $\varepsilon>0$ choose h_1 and h_2 as in Lemma 5·2. Then

$$F_i(w)=\int_0^{h_i(w)} e^{-\lambda t}f(x_t)dt, \qquad i=1, 2,$$

is bounded continuous on $\tilde{W}_{D_1}$ and hence by 5·2 $\mathrm{E}^x_{n,D_1}(F_i)\to\mathrm{E}^x_{D_1}(F_i)$, $i=1, 2$, as $n\to+\infty$. But, because

$$F_1(w)\le\int_0^{\sigma_D} e^{-\lambda t}f(x_t)dt\le F_2(w), \qquad \mathrm{E}^x_{n,D_1}(F_2-F_1)<\varepsilon\|f\|_{D_1},$$

one sees that $\lim u_n(x)=u(x)$. Q.E.D.

---

6) a is the center, and r is the radius of D.

Let E^d be the one-point compactification $R^d \cup \{\infty\}$ of R^d, W the space of continuous paths $w: t \in [0, +\infty) \to x(t, w) = x(t) = x_t \in E^d$ such that $x(t, w) \in R^d$ for $t < \tau_\infty$ and $= \infty$ for $t \geq \tau_\infty$, where $0 \leq \tau_\infty(w) \leq +\infty$, and **B** the corresponding coordinate Borel field. From Lemma 5·3, one can construct a family of probability measures $\{\mathbf{P}^x(\cdot), x \in R^d\}$ on **B** such that $[\mathrm{W}, \mathbf{B}, \mathbf{P}^x]$ is a diffusion and that $\mathbf{P}^x(x_t \in A, t < \sigma_D) = \mathbf{P}^x_D(x_t \in A)$, $x \in D$, $A \subset D$, for each open ball D (for example, by Bochner's projective limit construction of stochastic processes [1: p. 120] [11]). Denote by $\{T^t\}$ and $\{G^\lambda\}$ the corresponding semigroup and Green operators, and by $H_{\partial D}$ the harmonic measure:

$$H_{\partial D}(x, A) = \mathrm{E}^x[x(\sigma_D) \in A], \qquad x \in D, \quad A \subset \partial D.$$

LEMMA 5·4. i) *G^λ, $\lambda > 0$, maps $B(R^d)$ into $C(R^d)$, and if $f \in C(R^d)$, $u \equiv G^\lambda f$ is a solution of $(\lambda - \tilde{A})u = f$ in R^d.* ii) *$H_{\partial D}$ maps $B(\partial D)$ into $C(D)$.*

Proof. i) For each D, let $A_{n,D}$ be the approximating operators on D obtained by the method of 1.5 with φ (may depend on D) having sufficiently small support in $\{|x| < 1\}$. If $f \in C(R^d)$, then by the proof of Lemma 5·3,

$$u_{n,D}(x) = \mathrm{E}^x_{n,D}\left[\int_0^{+\infty} e^{-\lambda t} f(x_t) dt\right]$$

converges, as $n \uparrow +\infty$, to the function

$$u_D(x) = \mathrm{E}^x_D\left[\int_0^{\sigma_\infty} e^{-\lambda t} f(x_t) dt\right] = \mathrm{E}^x\left[\int_0^{\sigma_D} e^{-\lambda t} f(x_t) dt\right],$$

and, if $f \geq 0$, $u_D(x)$ increases to $u(x)$ as $D \uparrow R^d$. On the other hand, if in addition, $f \in C^1(R^d)$, $u_{n,D}$ is a bounded ($\|u_{n,D}\|_D \leq \lambda^{-1}\|f\|_D$) solution of $(\lambda - A_n)u = f$ in D, and hence applying the estimate in §1 to the quation $(\lambda - A_n)u_{n,D} = f$ in a fixed domain D_0 $\{u_{n,D}; n \geq 1, D \supset D_0, f \in C^1(R^d)$ and $\|f\|_{R^d} \leq 1\}$ turns out to be equicontinuous in D_0. Therefore $\{G^\lambda f; f \in C^1(R^d)$ and $\|f\|_{R^d} \leq 1\}$ is equicontinuous on each compact set, and hence the same for $\{G^\lambda f; f \in B(R^d)$ and $\|f\|_{R^d} \leq 1\}$. This implies that G^λ maps $B(R^d)$ into $C(R^d)$, and the rest is clear. The proof of ii) is much the same as above and is omitted.

We see easily that the range $G^\lambda(C(R^d))$ is independent of $\lambda > 0$, G^λ is one to one from $C(R^d)$ onto the common range and that $\lambda - (G^\lambda)^{-1}$ acting on the on the common range is also independent of $\lambda > 0$. $\lambda - (G^\lambda)^{-1}$ is called the generator of the diffusion $[\mathrm{W}, \mathrm{B}, \mathbf{P}^x]$. Summarizing the results in this section, we have

THEOREM 5·1. *There exists a diffusion* $[\mathbf{W}, \mathbf{B}, \mathbf{P}^x]$ *on* R^d *such that the Green opertor maps* $C(R^d)$ *into itself and the generator is a contraction of* $\bar{A}$.

If b^i is continuous ($1 \leq i \leq d$), and if f is a C^2-function with compact support, then f restricted on D belongs to $\mathfrak{D}(\bar{A}_D)$ provided D contains the support of f. Noting the uniform local character of the transition function of $\bar{A}$-diffusion, we have the following

THEOREM 5·2. *If* b^i *is continuous* ($1 \leq i \leq d$), *there exists a diffusion on* R^d *such that the corresponding semigroup* T^t *(acting on the space* $B(R^d)$*) satisfies*

$$\lim_{t \downarrow 0} \left\| \frac{T^t f - f}{t} - \sum_{i,j=1}^{d} a^{ij} \frac{\partial^2 f}{\partial x^i \partial x^j} - \sum_{i=1}^{d} b^i \frac{\partial f}{\partial x^i} \right\|_{R^d} = 0$$

for any C^2*-function* f *with compact support.*

References.

[1] S. BOCHNER, *Harmonic analysis and the theory of probability*, 1955.

[2] E. B. DYNKIN, *Theory of Markov processes* (English translation). Pergamon Press. 1961.

[3] I. B. GIRSANOV, *On stochastic integral equations of Ito.* D. A. H. CCCP. **138** (1961), 18-21.

[4] K. ITO, *On stochastic differential equations in a differentiable manifold.* Nagoya Math. J. **1** (1950), 35–47.

[5] S. ITO, *Fundamental solutions of parabolic differential equations and boundary value problems.* Jap. J. Math. 27 (1957), 55-102.

[6] C. B. MORREY, JR, *Second order elliptic system of differential equations.* Contribution to the theory of partial differential equations. Ann. Math. Studies. **33**, 101-159.

[7] E. NELSON, *An existence theorem for second order parabolic equations.* 88 (1958), 414-429.

[8] YU. V. PROHOROV, *Convergence of random processes and limit theorems in probability theory* (English translation). Theory of prob. and its appl. **1** (1956), 157–214.

[9] D. RAY, *Stationary Markov processes with continuous paths.* 82 (1956), 452-493.

[10] A. B. SKOROHOD, Исследования по теории случайных процессов, (1961).

[11] H. TOTOKI, *A Method of construction of measures on function spaces and its applications to stochastic processes.* Mem. Fac. Sci., Kyushu Univ. A. **15**, (1962), 178-190.

[12] H. TROTTER, *Approximation of semigroups of operators.* Pacific J. Math. 8 (1958), 887-919.

[13] A. D. VENTSEL, *On lateral conditions for multidimensional diffusion processes.*

Propagation of chaos for certain purely discontinuous Markov processes with interactions

Dedicated to Professor Kôsaku Yosida on his 60th birthday

By Hiroshi TANAKA

§ 1. Introduction.

Given a measurable space $(Q, \mathfrak{F})$ such that each single point set is in $\mathfrak{F}$, we consider a nonlinear equation for probability measures on Q:

$$\frac{du(t)}{dt}=\sum_{n=1}^{\infty}\int_Q q_n(x)\{\Pi_n(x, u(t), \cdot)-\delta(x, \cdot)\}u(t, dx) , \tag{1.1}$$

with the initial data f in $\mathfrak{M}$, the set of all probability measures on $(Q, \mathfrak{F})$. We assume that each $q_n(x)$ is a nonnegative $\mathfrak{F}$-measurable function on Q, $q(x)\equiv \sum_{n=1}^{\infty} q_n(x)<\infty$ and

$$\Pi_n(x,f,\Gamma)=\int\cdots\int\Pi_n(x, x_1, \cdots, x_n, \Gamma)f(dx_1)\cdots f(dx_n) , \quad (f\in\mathfrak{M}, \Gamma\in\mathfrak{F}) \tag{1.2}$$

where (i) $\Pi_n(x, x_1, \cdots, x_n, \cdot)\in\mathfrak{M}$ for each fixed $n, x_1, \cdots, x_n$, and (ii) $\Pi_n(x, x_1, \cdots, x_n, \Gamma)$ is measurable in $(x, x_1, \cdots, x_n)$ for each $\Gamma\in\mathfrak{F}$ and symmetric in $(x_1, \cdots, x_n)$ for each x and Γ. The precise meaning that $u(t)$ is a solution of (1.1) with $u(0)=f$ in $\mathfrak{M}$ is understood as follows. $u(t)$ is in $\mathfrak{M}$ for each $t\geq 0$ and satisfies

$$u(t, \Gamma)=\int_\Gamma e^{-q(x)t}f(dx) + \int_0^t ds\int_Q u(s, dy)\sum_{n=1}^{\infty}q_n(y)\int_\Gamma e^{-(t-s)q(z)}\Pi_n(y, u(s), dz),\ \Gamma\in\mathfrak{F},\ t\geq 0 . \tag{1.3}$$

The usual method of successive approximation provides us with the smallest solution of (1.3) among substochastic solutions, but of course, this smallest solution is not necessarily stochastic. We denote by $\mathfrak{U}$ the set of those $f\in\mathfrak{M}$ for which the smallest solutions of (1.3) belong to $\mathfrak{M}$ for each $t\geq 0$; it can happen that the class $\mathfrak{U}$ becomes empty.

Probabilistic significance is that the equation (1.1) corresponds to a Markov process describing the motion of one particle under the interactions with infinite number of similar particles. This class of Markov processes was first introduced and the explanation was given in the works of McKean [2] [3] through the so-called propagation of chaos, the notion which Kac [1] discovered for his 1-dimensional model of Maxwellian gas. In this direction, there are also works of D. P. Johnson [4] and T. Ueno [8], in which the propagation of chaos for the equation (1.1) were proved under a certain growth condition on $\{q_n(x)\}_{n\geq 1}$ when each $q_n(x)$ is bounded. The purpose of this paper is to prove that the propagation of chaos for (1.1) holds in the sense described below if the initial distribution f is taken from the class $\mathfrak{U}$. The method is similar to that of H. Tanaka [7] in its outline.

For each $n\geq 2$, we consider a Markov process $X(t)$ of n particles under motion. If each $q_k(x)$ is bounded, it is the Markov process on the state space $Q^n=Q\times \cdots \times Q$ with generator A_n given by

$$(A_n\varphi)(x_1, \cdots, x_n)=\sum_{N=1}^{n-1} n^{-N} \sum{}' q_n(x_i)\{\varphi_N(x_i, x_{i_1}, \cdots, x_{i_N})-\varphi\}\ , \tag{1.4}$$

where $\varphi_N(x_i, x_{i_1}, \cdots, x_{i_N})$ denotes the integral of $\varphi(\cdots, x_{i-1}, x, x_{i+1}, \cdots)$ viewed as a function of x alone with respect to the measure $\Pi_N(x_i, x_{i_1}, \cdots, x_{i_N}, dx)$, and $\sum'$ means the summation over all ordered $(N+1)$-tuples $(i, i_1, \cdots, i_N)$ of different integers taken from $\{1, \cdots, n\}$. But, if $q_k(x)$ is unbounded, the minimal Markov process corresponding to (1.4) is not necessarily conservative, that is, it can happen that the life time ζ is finite with positive probability. If this is the case, possibly some particles have jumped only finite number of times before ζ, and in our definition of the process $X(t)$ each of such particles continues the motion without dying until it performs infinite number of jumps. The precise definition of the process $X(t)$ will be given in § 4. Now our main result is stated as follows. For each $f\in\mathfrak{U}$ and $1\leq m<\infty$, the probability distribution at time t of m particles with arbitrary fixed coordinate numbers among n particles starting from the initial distribution $f^n=f\otimes \cdots \otimes f$ and performing the Markovian motion $X(\cdot)$ tends as $n\uparrow\infty$ to the m-fold outer product $u(t)^m=u(t)\otimes \cdots \otimes u(t)$ of the solution $u(t)$ of (1.3).

§ 2. Preliminaries.

In this paper we shall be concerned with several spaces such as the n-fold product space Q^n, the direct sum spaces $\boldsymbol{Q}_n=Q^1+ \cdots +Q^n$ and $\boldsymbol{Q}=Q^1+Q^2+ \cdots$; since Q is a measurable space, these spaces can also be considered as measurable spaces by the usual way. All functions we consider in this paper are real-valued and measurable, and so the words *real-valued* and *measurable* are often omitted for functions; the latter omission is the same for sets. The supremum norm of a bounded function φ is denoted by $\|\varphi\|$, and the integral over a set Γ of a function φ with respect to a (signed) measure f is denoted by $\langle f, \varphi\rangle_\Gamma$ or simply by $\langle f, \varphi\rangle$ if Γ is the whole space.

1°. *$\{q, \Pi\}$-minimal transition function.* Given a measurable space $(R, \mathfrak{S})$, we consider a kernel $A(x, \Gamma)$ over $(R, \mathfrak{S})$ expressed in the form:

$$A(x, \Gamma)=q(x)\{\Pi(x, \Gamma)-\delta(x, \Gamma)\}, \qquad x\in R,\ \Gamma\in\mathfrak{S},$$

where $0\le q(x)<\infty$, and $\Pi(x, \Gamma)$ is a substochastic measure on R for each x and $\mathfrak{S}$-measurable in x for each Γ. The minimal solution $p(t, x, \Gamma)$ of the backward equation associated with $A(x, \Gamma)$:

$$u(t, x, \Gamma)=e^{-q(x)t}\delta(x, \Gamma)+\int_0^t ds\int_R q(x)e^{-q(x)s}u(t-s, y, \Gamma)\Pi(x, dy) \tag{2.1}$$

is obtained by the usual successive approximation as follows. Set

$$p_0(t, x, \Gamma)=e^{-q(x)t}\delta(x, \Gamma) \tag{2.2a}$$

$$p_{k+1}(t, x, \Gamma)=p_0(t, x, \Gamma)+\int_0^t ds\int_R q(x)e^{-q(x)s}p_k(t-s, y, \Gamma)\Pi(x, dy)\ . \tag{2.2b}$$

Then, $p_k(t, x, \Gamma)$ increases to $p(t, x, \Gamma)$ as $k\uparrow\infty$; the p_k's are also obtained by

$$p_0(t, x, \Gamma)=e^{-q(x)t}\delta(x, \Gamma) \tag{2.3a}$$

$$p_{k+1}(t, x, \Gamma)=p_0(t, x, \Gamma)+\int_0^t ds\int_R p_k(s, x, dy)q(y)\int_\Gamma e^{-q(z)(t-s)}\Pi(y, dz)\ , \tag{2.3b}$$

and hence $p(t, x, \Gamma)$ is the minimal solution of the forward equation:

$$u(t, x, \Gamma)=e^{-q(x)t}\delta(x, \Gamma)+\int_0^t ds\int_R u(s, x, dy)q(y)\int_\Gamma e^{-q(z)(t-s)}\Pi(y, dz)\ . \tag{2.4}$$

$p(t, x, \Gamma)$ is called the $\{q, \Pi\}$-minimal transition function; it is a substochastic measure for each t and x.

We now prepare a simple convergence lemma for a sequence of $\{q_n, \Pi_n\}$-minimal transition functions. Let $0\le q_n(x)<\infty$ and $\Pi_n(x, \Gamma)=r_n(x)\Pi(x, \Gamma)$, and assume that (i) $\Pi(x, \cdot)$ and $\Pi_n(x, \cdot)$ are substochastic measures for each x and (ii) $q_n(x)$ and $r_n(x)$ converge respectively to finite limits $q(x)$ and $r(x)\equiv 1$ as $n\uparrow\infty$. Define $p_{k.n}(t, x, \Gamma)$ by (2.2) with q and Π replaced by q_n and Π_n respectively, and let $p_{.n}(t, x, \Gamma)$ be the $\{q_n, \Pi_n\}$-minimal transition function. Then, $p_{k.n}(t, x, \Gamma)$ *tends to* $p_k(t, x, \Gamma)$ *as* $n\uparrow\infty$ *for each* t, x *and* k, *and therefore if at a point* x *there exists an integer* $m\ge 1$ *such that*

$$p_{k,n}(t, x, \cdot)=p_{m,n}(t, x, \cdot) \quad \textit{for all } k\geq m \textit{ and } n\geq 1,$$

then $p_{\cdot,n}(t, x, \cdot)$ *tends to* $p(t, x, \cdot)$ *as* $n\uparrow\infty$.

2°. *A formula solving the equation* (1.3). S. Tanaka [7] and T. Ueno [9] expressed the (minimal) solution of the equation (1.3) in an explicit form; a part of their results is sketched here as a preparation for the next section. First we introduce the set $T=\bigcup_{n=1}^{\infty} T_n$, each set T_n being defined inductively as follows. (a) T_1 consists of single element θ. (b) If T_m is defined for $m<n$, then T_n is defined as the set consisting of all objects τ's of the form $\tau=[\tau_0, \tau_1, \cdots, \tau_k]$ where $\tau_j\in T_{n_j}$, $n_0+n_1\cdots+n_k=n$, $k\geq 1$. Next, we introduce the following notation for measure-valued t-functions $f_0(t), \cdots, f_k(t)$.

$$[f_0, f_1, \cdots, f_k](t, \Gamma)=\int_0^t ds\int_Q f_0(s, dx)q_k(x)\int_{Q^k} f_1(s, dx_1) \cdots f_k(s, dx_k)\int_\Gamma e^{-q(y)(t-s)}\Pi_k(x, x_1, \cdots, x_k, dy).$$

Given the equation (1.3), we now define the measure-valued t-function $u_\tau=u_\tau(t)$ indexed by $\tau\in T$ inductively as follows.

(i) $u_\theta=u_\theta(t, \Gamma)=\langle f, e^{-q(\cdot)t}\rangle_\Gamma$.

(ii) $u_\tau=[u_{\tau_0}, u_{\tau_1}, \cdots, u_{\tau_k}]$ for $\tau=[\tau_0, \tau_1, \cdots, \tau_k]$.

Then the (minimal) solution $u(t)$ of (1.3) is expressed as

$$u(t)=\sum_{\tau\in T} u_\tau(t)=\sum_{n=1} u_n(t)\ ;$$

here $u_n(t)=\sum_{\tau\in T_n} u_\tau(t)$ is defined inductively by

$$u_n=\sum_{k=1}^{n-1}\sum_{\substack{n_0+\cdots+n_k=n\\ n_0,\cdots,n_k\geq 1}}[u_{n_0}, u_{n_1}, \cdots, u_{n_k}], \qquad u_1=u_\theta\ . \tag{2.5}$$

§ 3. A formula concerning the minimal solution of (1.3).

In this section the formula given in the second half of the preceding section is discussed from another point of view; the resulting formula (3.2) will be the basis for the proof of the propagation of chaos.

Let φ_m be a function on Q^m. Writing $\Pi_N(x_i, x_{i_1}, \cdots, x_{i_N}, \varphi_m)$ for the integral

$$\int_Q \Pi_N(x_i, x_{i_1}, \cdots, x_{i_N}, dx)\varphi_m(\cdots, x_{i-1}, x, x_{i+1}, \cdots) \ (i_1, \cdots, i_N>m)$$

and regarding it as a functions of $\max_{1\leq k\leq N} i_k$ variables, we collect the functions of k variables in the formal series:

$$\sum_{m=1}^{\infty}\sum_{N=1}^{\infty}\sum_{i=1}^{m} q_N(x_i)\{\Pi_N(x_i, x_{m+1}, \cdots, x_{m+N}, \varphi_m)-\varphi_m\} .$$

The result is

$$\sum_{m=1}^{k-1}\sum_{i=1}^{m} q_{k-m}(x_i)\Pi_{k-m}(x_i, x_{m+1}, \cdots, x_k, \varphi_m)-\sum_{i=1}^{k} q(x_i)\varphi_m$$

for $k\geq 2$ and $-q(x_1)\varphi_1(x_1)$ for $k=1$. On the basis of this result, we introduce a $\{q, \Pi\}$-function on $\boldsymbol{Q}$ as follows. For $\boldsymbol{x}=(x_1, \cdots, x_k)\in Q^k$

$$\boldsymbol{q}(\boldsymbol{x})=q(x_1)+\cdots+q(x_k) , \tag{3.1a}$$

$$\begin{aligned}\boldsymbol{\Pi}(x, \Gamma)&=\boldsymbol{q}(\boldsymbol{x})^{-1}\sum_{i=1}^{m} q_{k-m}(x_i)\Pi_{k-m}(x_i, x_{m+1}, \cdots, x_k, \chi_\Gamma)\\ &\qquad \text{for } \Gamma\subset Q^m \text{ with } 1\leq m\leq k-1\\ &=0 \quad \text{for } \Gamma\subset Q^m \text{ with } m\geq k\\ &=0 \quad \text{for all } \Gamma \text{ if } k=1 .\end{aligned} \tag{3.1b}$$

An easy calculation shows that $\boldsymbol{\Pi}$ is substochastic, and so the $\{\boldsymbol{q}, \boldsymbol{\Pi}\}$-minimal transition function $P(t, x, \cdot)$ is substochastic.

Denote by Φ_k the space of bounded measurable functions on Q^k, and by Φ the space of those functions on $\boldsymbol{Q}$ whose restrictions on each Q^k belong to Φ_k. The restriction of $\boldsymbol{\varphi}\in\Phi$ on Q^k is denoted by φ_k or $(\boldsymbol{\varphi})_k$. Then Φ is a Fréchet space with seminorms $\|\boldsymbol{\varphi}\|_k=\sum_{1\leq j\leq k}\|\varphi_j\|$. The transition function $P(t, x, \cdot)$ induces in the usual way a semigroup $\{T^t\}$ on Φ which is positive and $T^t 1\leq 1$. Our task in this section is to prove the following theorem.

THEOREM 1. *For any* $\varphi\in\Phi_k$,

$$\langle u(t)^k, \varphi\rangle=\sum_{m=k}^{\infty}\langle f^m, (T^t\boldsymbol{\varphi})_m\rangle \tag{3.2}$$

where $u(t)$ *is the minimal solution of* (1.3), *and* $(\boldsymbol{\varphi})_m=\varphi$ (*for* $m=k$), $=0$ (*for* $m\neq k$).

Before giving the proof of this theorem, we introduce some notations. Denote by Ξ_k the space of functions ξ's on $\mathfrak{M}$ which are expressed as $\xi(f)=\langle f^k, \varphi\rangle$ for some $\varphi\in\Phi_k$; here φ can be chosen to be symmetric. By setting $\|\xi\|_k=\|\varphi\|$ for $\xi(f)=\langle f^k, \varphi\rangle$ with symmetric $\varphi\in\Phi_k$, Ξ_k becomes a Banach space with the norm $\|\xi\|_k$, since a bounded symmetric function φ on Q^k is recovered by the value $\xi(f)=\langle f^k, \varphi\rangle$ for f running on $\mathfrak{M}$ as the following formula indicates:

$$\varphi(x_1, \cdots, x_k)=\frac{1}{k!}\sum_{m=1}^{k}(-1)^{k-m}\sum_{J\in\mathfrak{S}_m} m^k\langle f_J^k, \varphi\rangle , \tag{3.3}$$

where $\mathfrak{J}_m$ is the family of all subsets of $\{1, \cdots, k\}$ with m elements, and f_J^k denotes the k-fold outer product of f_J defined by $f_J(\Gamma)=\frac{1}{m}\sum_{j\in J}\delta(x_j, \Gamma)$ for $J\in\mathfrak{J}_m$ (see [6]). Let Ξ be the space of all sequences $\boldsymbol{\xi}=(\xi_1, \xi_2, \cdots)$ with each ξ_j in Ξ_j, and define a multiplication $\boldsymbol{\xi}\otimes\boldsymbol{\eta}$ by $(\boldsymbol{\xi}\otimes\boldsymbol{\eta})_k=\sum_{i+j=k}\xi_i\eta_j$. Then, Ξ is a Fréchet space with the seminorms $\|\boldsymbol{\xi}\|_k=\sum_{1\le j\le k}\|\xi_j\|_j$, and $\|\boldsymbol{\xi}\otimes\boldsymbol{\eta}\|_k\le\|\boldsymbol{\xi}\|_k\|\boldsymbol{\eta}\|_k$ holds for all k. The mapping $\mathfrak{S}:\Phi\to\Xi$ defined by

$$(\mathfrak{S}\boldsymbol{\varphi})_k(f)=\langle f^k, \varphi_k\rangle \quad \text{for} \quad \boldsymbol{\varphi}=(\varphi_k)_{k\ge 1}$$

satisfies $\|\mathfrak{S}\boldsymbol{\varphi}\|_k\le\|\boldsymbol{\varphi}\|_k$ for all k.

Now the proof of Theorem 1 is carried out in 3 steps.

Step 1. Suppose $q(x)$ is bounded. In this case, the linear operator $A:\Phi\to\Phi$ defined by

$$(A\boldsymbol{\varphi})_k(x_1, \cdots, x_k)=\sum_{m=1}^{k-1}\sum_{i=1}^{m} q_{k-m}(x_i)\Pi_{k-m}(x_i, x_{m+1}, \cdots, x_k, \varphi_m)$$
$$-\sum_{m=1}^{k} q(x_m)\varphi_k(x_1, \cdots, x_k)$$

(the first term vanishes for $k=1$)

is bounded, and the semigroup with generator A is nothing but $\{T^t\}$. Since one can easily prove that $\langle f^m, \varphi\rangle=0$ for all $f\in\mathfrak{M}$ implies

$$\int\cdots\int\sum_{i=1}^{m} q_{k-m}(x_i)\Pi_{k-m}(x_i, x_{m+1}, \cdots, x_k, \varphi)f(dx_1)\cdots f(dx_k)=0 ,$$

one can define a linear operator $\mathfrak{A}:\Xi\to\Xi$ by $\mathfrak{A}\mathfrak{S}\boldsymbol{\varphi}=\mathfrak{S}A\boldsymbol{\varphi}$. Then obviously the semigroup $\{\mathfrak{T}^t\}$ with generator $\mathfrak{A}$ satisfies $\mathfrak{T}^t\mathfrak{S}\boldsymbol{\varphi}=\mathfrak{S}T^t\boldsymbol{\varphi}$. Moreover, by direct computations one can show that $\mathfrak{A}$ is a derivation of the algebra Ξ, that is $\mathfrak{A}(\boldsymbol{\xi}\otimes\boldsymbol{\eta})=(\mathfrak{A}\boldsymbol{\xi})\otimes\boldsymbol{\eta}+\boldsymbol{\xi}\otimes(\mathfrak{A}\boldsymbol{\eta})$, and this fact is reflected to the multiplicative property of $\{\mathfrak{T}^t\}$:

$$\mathfrak{T}^t(\boldsymbol{\xi}\otimes\boldsymbol{\eta})=(\mathfrak{T}^t\boldsymbol{\xi})\otimes(\mathfrak{T}^t\boldsymbol{\eta}) . \tag{3.4}$$

Step 2. If $q(x)$ is unbounded, we set $q_k^{(n)}(x)=q_k(x)\wedge n$ (for $k\le n$), $=0$ (for $k>n$) and $q^{(n)}(x)=\sum_{k=1}^{n} q_k^{(n)}(x)$. Let $P_n(t, x, \Gamma)$ be the $\{\boldsymbol{q}_n, \boldsymbol{\Pi}_n\}$-minimal transition function and $\{\mathfrak{T}_n^t\}$ the associated semigroup on Ξ, where $\boldsymbol{q}_n$ and $\boldsymbol{\Pi}_n$ are defined by (3.1) with $q_k(x)$ and $q(x)$ replaced by $q_k^{(n)}(x)$ and $q^{(n)}(x)$ respectively. Since $q^{(n)}(x)$ is bounded for each n, the semigroup $\{\mathfrak{T}_n^t\}$ is multiplicative by Step 1. On the other hand, by the convergence lemma mentioned in 1° of §2, the transition function $P_n(t, x, \Gamma)$ tends to $P(t, x, \Gamma)$ as $n\uparrow\infty$ for each $\Gamma\subset Q^k$ and hence

$$(\mathfrak{S}T^t\boldsymbol{\varphi})_k(f)=\langle f^k, \sum_{j\le k}\int_{Q^j}P(t,\cdot,d\boldsymbol{y})\varphi_j(\boldsymbol{y})\rangle_{Q^k}$$
$$=\lim_{n\to\infty}\langle f^k, \sum_{j\le k}\int_{Q^j}P_n(t,\cdot,d\boldsymbol{y})\varphi_j(\boldsymbol{y})\rangle_{Q^k}$$
$$=\lim_{n\to\infty}(\mathfrak{T}_n^t\mathfrak{S}\boldsymbol{\varphi})_k(f)\ .$$

Therefore, one can define $\mathfrak{T}^t$ on Ξ by $\mathfrak{T}^t\mathfrak{S}\boldsymbol{\varphi}=\mathfrak{S}T^t\boldsymbol{\varphi}$ and so-defined $\{\mathfrak{T}^t\}$ becomes a *multiplicative* semigroup as the limit of such ones.

Step 3. For each $t\ge 0$ and integers k, n with $1\le k\le n$ we define a measure $f_{n,k}(t)$ on Q^k by $f_{n,k}(t,\Gamma)=\langle f^n, P(t,\cdot,\Gamma)\rangle_{Q^n}$, $\Gamma\subset Q^k$. If $\varphi(x_1,\cdots,x_k)=\varphi_1(x_1)\cdots\varphi_k(x_k)$, then

$$\langle f_{n,k}(t),\varphi\rangle=(\mathfrak{T}^t\boldsymbol{\xi}_1\otimes\cdots\otimes\boldsymbol{\xi}_k)_n(f)$$

where $\boldsymbol{\xi}_j(f)=(\langle f,\varphi_j\rangle, 0, 0,\cdots)$, and hence by the multiplicative property (3.4) of $\{\mathfrak{T}^t\}$ we have

$$f_{n,k}(t)=\sum_{n_1+\cdots+n_k=n}f_{n_1,1}(t)\otimes\cdots\otimes f_{n_k,1}(t)\ . \tag{3.5}$$

On the other hand, by the forward equation for $P(t,\cdot,\cdot)$ we have for $\Gamma\subset Q^1$

$$P(t,x,\Gamma)=e^{-q(x_1)t}\delta(x_1,\Gamma)\qquad\text{if}\quad x\in Q^1$$
$$=\int_0^t ds\sum_{m=2}^{n}\int_{Q^m}P(s,x,dy)q_{m-1}(y_1)\int_\Gamma e^{-q(z)(t-s)}\Pi_{m-1}(y_1,y_2,\cdots,y_m,dz)$$
$$\text{if}\quad x\in Q^n,\ n\ge 2\ .$$

Integrating both sides of the above by f^n and then using (3.5),

$$f_{1,1}(t,\Gamma)=\langle f, e^{-q(\cdot)t}\delta(\cdot,\Gamma)\rangle=u_\theta(t,\Gamma)$$
$$f_{n,1}(t,\Gamma)=\int_0^t ds\sum_{m=2}^{n}\int_{Q^m}f_{n,m}(s,d\boldsymbol{y})q_{m-1}(y_1)\int_\Gamma e^{-q(z)(t-s)}\Pi_{m-1}(y_1,y_2,\cdots,y_m,dz)$$
$$=\int_0^t ds\sum_{m=2}^{n}\sum_{n_1+\cdots+n_m=n}\int_{Q^m}f_{n_1,1}(s,dy_1)\cdots f_{n_m,1}(s,dy_m)q_{m-1}(y_1)$$
$$\times\int_\Gamma e^{-q(z)(t-s)}\Pi_{m-1}(y_1,y_2,\cdots,y_m,dz)$$
$$=\sum_{m=2}^{n}\sum_{n_1+\cdots+n_m=n}[f_{n_1,1},\cdots,f_{n_m,1}]\qquad\text{for}\quad n\ge 2\ .$$

Comparing this with (2.5), we conclude that $u(t)=\sum_{n=1}^{\infty}f_{n,1}(t)$, and this means that (3.2) holds for $k=1$. For $k\ge 2$, using (3.5) again, we have

$$\sum_{n=k}^{\infty}f_{n,k}(t)=\sum_{n=k}^{\infty}\sum_{n_1+\cdots+n_k=n}f_{n_1,1}(t)\otimes\cdots\otimes f_{n_k,1}(t)=(\sum_{n=1}^{\infty}f_{n,1}(t))^k=u(t)^k\ ,$$

completing the proof of the theorem.

§ 4. The motion of *n* particles.

Let $\bar{Q}$ be the adjoined space $Q \smile \{\Delta\}$, Δ being an extra point $\notin Q$. In this section let $n \geq 2$ be fixed. The motion of n particles we consider in this section is a Markov process $X(t)=(X_1(t), \cdots, X_n(t))$ with state space $\bar{Q}^n=\bar{Q}\times \cdots \times\bar{Q}$ whose probabilistic development is determined in the following way. Suppose the process starts at $\boldsymbol{x}=(x_1, \cdots, x_n)\in\bar{Q}^n$ and set $I=\{i: x_i\in Q\}$, $I'=\{i; x_i=\Delta\}$. Then,

(1)
$$X_i(t)=\Delta \quad \text{for} \quad i\in I' \quad \text{and all} \quad t\geq 0\,,$$
$$X_i(t)=x_i \quad \text{for} \quad i\in I,\ t<S_1\,,$$

where S_1 is a random variable with

$$\boldsymbol{P}_x\{S_1>t\}=\exp\{-t\sum_{i\in I} r(x_i)\} \qquad (t\geq 0)$$
$$r(x_i)=\sum_{N=1}^{n-1} n^{-N}\frac{(n-1)!}{(n-1-N)!}\,q_N(x_i)\,.$$

(2) At time S_1, one of the particles with index in I jumps according to the following probability law. For each $i\in I$

$$\boldsymbol{P}_x\{X_i(S_1)\in\Gamma,\ X_j(S_1)=x_j,\ j\neq i\}$$
$$=\sum_{N=1}^{n-1} n^{-N}\sum{}' q_N(x_i)\Pi_N(x_i, x_{i_1}, \cdots, x_{i_N}, \Gamma)/\sum_{j\in I} r(x_j),\ \Gamma\subset Q\,.$$

For $\Gamma=\{\Delta\}$, the left hand side of the above is replaced by

$$\{r(x_i)-\sum_{N=1}^{n-1} n^{-N}\sum{}' q_N(x_i)\Pi_N(x_i, x_{i_1}, \cdots, x_{i_N}, Q\}/\sum_{j\in I} r(x_j)\,.$$

Here, we set $\Pi_N(x_i, x_{i_1}, \cdots, x_{i_N}, \Gamma)=0$ for $\Gamma\subset Q$ if at least one of $x_i, x_{i_1}, \cdots, x_{i_N}$ is Δ, and $\sum'$ is taken over all ordered N-tuples $(i_1, \cdots, i_N)$ such that $i\leq i_1, \cdots, i_N\leq n$ and $i_1, \cdots, i_N$ $(\neq i)$ are all different.

(3) After the time S_1, we define the process $X(t)$ up to the second jumping time S_2 so that the process $X(S_1+t)$ conditional to the σ-field generated by $\{X_1(t_1\wedge S_1)\in\Gamma_1, \cdots, X_n(t_n\wedge S_1)\in\Gamma_n\}$ is exactly the same in law as the one defined by (1) and (2) with initial state $X(S_1)$. We proceed in the same way after the second jumping time S_2, the third jumping time S_3 and so on. Thus we obtain the process $X(t)$ for $t<S=\lim_{n\to\infty} S_n$.

(4) If $S(\omega)<\infty$, then each component process $X_i(t)$ has performed either infinitely many jumps or finitely many (including no) jumps. In the former case we set $X_i(S)=\Delta$ and in the latter case $X_i(S)=X_i(S-)$, and then start afresh

at time S as in (1) (2) and (3). Repeated arguments of this now leads to the process $X(t)$ defined for all $t \geq 0$.

The following explicit construction of the process $X(t)$ will be used later. For each $\boldsymbol{x}=(x_1, \cdots, x_n) \in \bar{Q}^n$ we set

$$\alpha(t, \boldsymbol{x})=\exp\{-t\sum_{i=1}^{n} r(x_i)\}, \qquad r(\Delta)=0$$

$$\beta(\boldsymbol{i}, \boldsymbol{x})=n^{-N}q_N(x_i)/\sum_{j=1}^{n} r(x_j), \ \boldsymbol{i}=(i, i_1, \cdots, i_N), \ q_N(\Delta)=0$$

$$\gamma(\boldsymbol{i}, \boldsymbol{x}, \Gamma)=\begin{cases} \Pi_N(x_i, x_{i_1}, \cdots, x_{i_N}, \Gamma), & \text{for } \Gamma \subset Q \\ 0 & \text{, for } \Gamma=\{\Delta\} \text{ if all } x_i, x_{i_1}, \cdots, x_{i_N} \text{ are in } Q \\ 1 & \text{, for } \Gamma=\{\Delta\} \text{ if at least one of } x_i, x_{i_1}, \cdots, x_{i_N} \text{ is } \Delta, \end{cases}$$

and then construct a discrete parameter Markov process $U_k=(R_k, Y_k, Z_k)$, $k \geq 1$, over a suitable probability space $(\Omega, \boldsymbol{P}_x)$ with the following properties.

(i)
$$0<R_k \leq \infty, \ Y_k \in \mathfrak{J}, \ Z_{k.m}(\boldsymbol{i}) \in \bar{Q},$$
$$Z_k=\{Z_k(\boldsymbol{i}), \boldsymbol{i} \in \mathfrak{J}\}=\{(Z_{k.1}(\boldsymbol{i}), \cdots, Z_{k.n}(\boldsymbol{i})), \boldsymbol{i} \in \mathfrak{J}\};$$

here $\mathfrak{J}$ denotes the set of all $\boldsymbol{i}=(i, i_1, \cdots, i_N)$ such that $i, i_1, \cdots, i_N$ are different integers taken from $1, \cdots, n$.

(ii) The initial distribution is given by

$$\begin{aligned} \boldsymbol{P}_x\{R_1>t, \ Y_1=\boldsymbol{i}, \ Z_{1.m}(\boldsymbol{l}) \in \Gamma_{m.\boldsymbol{l}}, 1 \leq m \leq n, \boldsymbol{l} \in \mathfrak{J}\} \\ =\alpha(t, \boldsymbol{x})\beta(\boldsymbol{i}, \boldsymbol{x}) \prod_{\substack{(m,\boldsymbol{l}) \\ m=|\boldsymbol{l}|}} \gamma(\boldsymbol{l}, x, \Gamma_{m.\boldsymbol{l}}) \prod_{\substack{(m,\boldsymbol{l}) \\ m \neq |\boldsymbol{l}|}} \delta(x_m, \Gamma_{m.\boldsymbol{l}}) \end{aligned} \tag{4.1}$$

where $|\boldsymbol{l}|=l$ for $\boldsymbol{l}=(l, l_1, \cdots, l_N)$.

(iii) For $\Lambda=\{R_{k+1}>t, \ Y_{k+1}=\boldsymbol{i}, \ Z_{k+1.m}(\boldsymbol{l}) \in \Gamma_{m.\boldsymbol{l}}, 1 \leq m \leq n, \boldsymbol{l} \in \mathfrak{J}\}$, the conditional probability $\boldsymbol{P}_x\{\Lambda | U_1, \cdots, U_k\}=\boldsymbol{P}_x\{\Lambda | U_k\}$ is given by the right hand side of (4.1) with $\boldsymbol{x}$ replaced by $Z_k(Y_k)$ almost surely on the set $\{\sum_{m=1}^{n} r(Z_{k.m}(Y_k))>0\}$, while we set $U_j=\Delta'$ for $j \geq k+1$ on the set $\{\sum_{m=1}^{n} r(Z_{k.m}(Y_k))=0\}$, Δ' being an extra point (death point).

In order to construct the process $X(t)$ by means of $\{U_k\}$, we need to extend the scope of the time parameter of $\{U_k\}$ further to

$$\infty+1, \infty+2, \cdots, 2\infty+1, 2\infty+2, \cdots, (n-1)\infty+1, (n-1)\infty+2, \cdots.$$

We set

$$(4.2) \qquad x_i(\infty)=\begin{cases} \Delta & \text{if } \#\{k: |Y_k|=i\}=\infty \\ Z_{k(i),i}(Y_{k(i)}) & \text{if } k(i)=\sup\{1\le k<\infty\,;\,|Y_k|=i\}<\infty\,, \end{cases}$$

and then construct $\{U_{\infty+k}\}_{k=1,2,\cdots}$ so that its distribution conditional to $\{U_1, U_2, \cdots\}$ coincides with that of the process obtained exactly by the same way as in (i) (ii) and (iii) but with the replacement of $\boldsymbol{x}$ by $\boldsymbol{x}(\infty)=(x_1(\infty), \cdots, x_n(\infty))$. Repeating this kind of construction again and again, we obtain a Markov process $\{U_k\}$ with time parameters

$$1, 2, \cdots, \infty+1, \infty+2, \cdots, 2\infty+1, 2\infty+2, \cdots, (n-1)\infty+1, (n-1)\infty+2, \cdots .$$

Now the process $X(t)$ is constructed over $(\Omega, \boldsymbol{P}_x)$ by

$$\begin{aligned} X(t)&=Z_{k-1}(Y_{k-1}) \quad \text{if} \quad \sum_{j<k} R_j\le t<\sum_{j\le k} R_j \text{ [1]} \\ &=(\Delta, \cdots, \Delta) \quad \text{if} \quad \sum_{j<n\infty} R_j\le t<+\infty\,. \end{aligned}$$

For a subset I of $\{1, \cdots, n\}$, we define the first jumping time $T(I)$ of the particles with index in I by

$$(4.3) \qquad \begin{aligned} T(I)&=\sum_{j\le k(I)} R_j \quad \text{if} \quad k(I)=\min\{k: |Y_k|\in I\}<n\infty \\ &=+\infty \qquad \text{if} \quad |Y_k|\notin I \quad \text{for all} \quad k<n\infty\,; \end{aligned}$$

it is distributed according to the exponential distribution with mean $1/\sum_{i\in I} r(x_i)$.

Next, we introduce a (q, Π)-function on the space $\boldsymbol{Q}_n$ as follows: for $\boldsymbol{x}=(x_1, \cdots, x_k)\in Q^k$

$$\begin{aligned} \boldsymbol{q}_n(\boldsymbol{x})&=\sum_{i=1}^{k} r(x_i) \\ \boldsymbol{\Pi}_n(x, \Gamma)&=\boldsymbol{q}_n(\boldsymbol{x})^{-1}n^{-(k-m)}\frac{(n-m)!}{(n-k)!}\sum_{i=1}^{m} q_{k-m}(x_i)\Pi_{k-m}(x_i, x_{m+1}, \cdots, x_k, \chi_\Gamma) \\ &\qquad\qquad \text{for } \Gamma\subset Q^m \text{ with } 1\le m\le k-1 \\ &=0 \qquad\qquad \text{for } \Gamma\subset Q^m \text{ with } k\le m\le n \\ &=0 \qquad\qquad \text{for all } \Gamma \text{ if } k=1\,, \end{aligned}$$

and denote by $P_n(t, \boldsymbol{x}, \Gamma)$ the $\{\boldsymbol{q}_n, \boldsymbol{\Pi}_n\}$-minimal transition function.

PROPOSITION 2. *If φ is a nonnegative bounded function on Q^m $(m\le n)$, then for any $f\in\mathfrak{M}$*

---

[1] We set $R_0=0$, $Z_0(Y_0)=\boldsymbol{x}$ and $Z_{k-1}(Y_{k-1})=\boldsymbol{x}(j\infty)$ if $k=j\infty+1$, where $\boldsymbol{x}(j\infty)$ is defined similarly to (4.2).

$$\sum_{k=m}^{n} \langle f^k, \int_{Q^m} P_n(t, \cdot, dy)\varphi(y)\rangle_{Q^k}$$
$$\leq \int_{Q^n} E_{(x_1,\dots,x_n)}\{\varphi(X_1(t), \cdots, X_m(t))\} f(dx_1) \cdots f(dx_n)^{2)} .$$

PROOF. Taking a family $\sigma=\{I_0, \cdots, I_p\}$ of subsets of $\{1, \cdots, n\}$ such that

$$I_0=\{1, \cdots, m\}, \quad I_{j-1}\subsetneqq I_j \quad (1\leq j\leq p) \tag{4.4}$$

and recalling (4.3), we set inductively

$$T(I_p, I_{p-1}, \cdots, I_{p'})= \sum_{j\leq k(I_p, I_{p-1}, \cdots, I_{p'})} R_j$$

$$\text{if} \quad k(I_p, \cdots, I_{p'})=\min\{k>k(I_p, \cdots, I_{p'+1}): |Y_k|\in I_{p'}\}<n\infty ,$$

$$T(I_p, I_{p-1}, \cdots, I_{p'})=+\infty$$

$$\text{if} \quad T(I_p, \cdots, I_{p'+1})=+\infty \text{ or if } |Y_k|\notin I_{p'} \text{ for } k(I_p, \cdots, I_{p'+1})<k<n\infty ,$$

and then

$$\Lambda_\sigma=\{T(I_p)\leq t,\ Y_{k(I_p)}\in \tilde{I}_p,\ T(I_p, I_{p-1})\leq t,\ Y_{k(I_p, I_{p-1})}\in \tilde{I}_{p-1}$$
$$\cdots,\ T(I_p, \cdots, I_1)\leq t,\ Y_{k(I_p, \cdots, I_1)}\in \tilde{I}_1,\ T(I_p, \cdots, I_0)>t\} ,$$

where $\tilde{I}_j$ denotes the set of all ordered $(p'+1)$-tuples $(i, i_1, \cdots, i_{p'})$ such that $i\in I_{j-1}$ and $\{i_1, \cdots, i_{p'}\}=I_j-I_{j-1}$ $(p'=\#(I_j-I_{j-1}))$. If we set

$$\varphi(t, \sigma)=\int_{\Lambda_\sigma} \varphi(X_1(t), \cdots, X_m(t))P_{(x_1, \cdots, x_n)}(d\omega) ,$$

then $\varphi(t, \sigma)$ is a function of $x(I_p)=(x_j, j\in I_p)$ alone for each t and σ, and

$$\varphi(t, \sigma)=\int_0^t e^{-sr(I_p)}n^{-N}N! \sum_{i\in I_p} q_N(x_i)\Pi_N(x_i, x(I_p-I_{p-1}), \varphi(t-s, \sigma_1))ds \tag{4.5}$$

where $\qquad N=\#(I_p-I_{p-1}),\ \sigma_1=\{I_0, \cdots, I_{p-1}\}$ and $r(I_p)=\sum_{i\in I_p} r(x_i)$.

Two $\sigma=\{I_0, \cdots, I_p\}$ and $\sigma'=\{I_0, I_1', \cdots, I_{p'}'\}$ both satisfying (4.4) are said to be similar, if $p=p'$ and $\# I_j=\# I_j'$ for $1\leq j\leq p$. Using (4.5) we can easily prove by induction on p that $\langle f^n, \varphi(t, \sigma)\rangle=\langle f^n, \varphi(t, \sigma')\rangle$ for similar σ and σ', and so, if σ is of the form:

$$\{\{1, 2, \cdots, m\}, \{1, 2, \cdots, l\}, \cdots, \{1, 2, \cdots, j\}, \{1, 2, \cdots, k\}\} \tag{4.6}$$
$$(m<l< \cdots <j<k)$$

then the function $\tilde{\varphi}(t, \sigma)=$(the number of σ' similar to σ)$\varphi(t, \sigma)$ satisfies

2) We set $\varphi(x_1, \cdots, x_m)=0$ if at least one of $x_1, \cdots, x_m$ is Δ.

$$(4.7)\qquad \langle f^n, \tilde{\varphi}(t,\sigma)\rangle = \langle f^n, \sum_{\sigma': \text{ similar to } \sigma} \varphi(t,\sigma')\rangle$$

$$(4.8a)\qquad \tilde{\varphi}(t,\Gamma) = \int_0^t e^{-sr(\Gamma_p)} n^{-(k-j)} \frac{(n-j)!}{(n-k)!} \sum_{i=1}^{j} q_{k-j}(x_i) \Pi_N(x_i, x_{j+1}, \cdots, x_k, \tilde{\varphi}(t-s,\sigma_1))ds$$

$$(\sigma_1 = \{\{1,\cdots,m\},\{1,\cdots,l\},\cdots,\{1,\cdots,j\}\}, p>1)$$

$$(4.8b)\qquad \tilde{\varphi}(t,\sigma) = e^{-tr(\Gamma_0)}\varphi(x_1,\cdots,x_m),\ \sigma=\{\{1,\cdots,m\}\}\ .$$

Therefore, the function $\varphi_k(t,(x_1,\cdots,x_k))$ defined by

$$\varphi_k(t,(x_1,\cdots,x_k)) = \begin{cases} \displaystyle\sum_{\substack{\sigma \text{ is of the form (4.6)} \\ \text{with fixed } k}} \tilde{\varphi}(t,\sigma), & \text{for } m<k \\ \tilde{\varphi}(t,\sigma), \text{ for } m=k & (\sigma=\{\{1,\cdots,m\}\}\ , \end{cases}$$

satisfies

$$(4.9)\qquad \varphi_k(t,(x_1,\cdots,x_k)) = \int_0^t e^{-sq_n(x_1,\cdots,x_k)} \sum_{j=1}^{k-1} n^{-(k-j)} \frac{(n-j)!}{(n-k)!}$$

$$\times \sum_{i=1}^{j} q_{k-j}(x_i) \Pi_N(x_i, x_{j+1},\cdots,x_k,\varphi_j(t-s))ds,\ k>m$$

which is nothing but the backward equation associated with $\{q_n, \Pi_n\}$, and hence $\varphi_k(t,(x_1,\cdots,x_k))$ coincides with $\langle P_n(t,(x_1,\cdots,x_k),\cdot),\varphi\rangle_{Q^m}$. On the other hand, by the definition of Λ_σ two events Λ_σ and $\Lambda_{\sigma'}$ are mutually exclusive unless σ and σ' are the same, and so we have

$$\int_{Q^n} \boldsymbol{E}_{(x_1,\cdots,x_n)}\{\varphi(X_1(t),\cdots,X_m(t))\} f(dx_1)\cdots f(dx_n)$$

$$\geq \sum_\sigma \langle f^n, \varphi(t,\sigma)\rangle = \langle f^n, \sum_{\sigma \text{ is of the form (4.6)}} \tilde{\varphi}(t,\sigma)\rangle$$

$$= \sum_{k=m}^{n} \langle f^k, \varphi_k(t)\rangle = \sum_{k=m}^{n} \langle f^k, \int_{Q^m} P_n(t,\cdot,dy)\varphi(y)\rangle\ ,$$

completing the proof of the proposition.

§ 5. Propagation of chaos.

In the preceding section we fixed n, the number of particles under motion, but here we let $n\uparrow\infty$ and prove that the propagation of chaos holds if the initial distribution f of (1.3) belongs to the class $\mathfrak{U}$. Let $u_n(t)$ be the probability distribution at time t of n particles under motion starting with the initial distribution f^n, that is

$$u_n(t,\Gamma) = \int_{Q^n} \boldsymbol{P}_{(x_1,\cdots,x_n)}\{X(t)\in\Gamma\} f(dx_1)\cdots f(dx_n)\ ,$$

where $\{X(t), \boldsymbol{P}_{(x_1, \cdots, x_n)}\}$ is the Markov process on $\bar{Q}^n$ introduced in § 4.

THEOREM 3. *For each* $m \geq 1$, $f \in \mathfrak{U}$ *and a bounded function* φ *on* Q^m,

$$\int_{Q^m \times \bar{Q}^{n-m}} \varphi(x_1, \cdots, x_m) u_n(t, dx_1, \cdots, dx_n) \to \langle u(t)^m, \varphi \rangle \text{ as } n \uparrow \infty , \tag{5.1}$$

where $u(t)$ *is the solution of* (1.3).

PROOF. By the convergence lemma in 1° of § 2, $P_n(t, \boldsymbol{x}, \Gamma)$ tends to $P(t, \boldsymbol{x}, \Gamma)$ as $n \uparrow \infty$ for each t, $\boldsymbol{x}$ and $\Gamma \subset Q^m$, and so if we set for $\Gamma \subset Q^m$

$$F_{n,k}(t, \Gamma) = \langle f^k, P_n(t, \cdot, \Gamma) \rangle_{Q^k} , \quad F_k(t, \Gamma) = \langle f^k, P(t, \cdot, \Gamma) \rangle_{Q^k} ,$$

then $F_{n,k}(t, \Gamma)$ tends to $F_k(t, \Gamma)$ as $n \uparrow \infty$. Therefore, by Theorem 1 and Proposition 2 and using Fatou's lemma, we have

$$1 \geq \varliminf_{n\to\infty} \sum_{k=1}^{n} F_{n,k}(t, \Gamma) \geq \sum_{k=1}^{\infty} F_k(t, \Gamma) = u(t)^m(\Gamma) .$$

Since $f \in \mathfrak{U}$ implies $u(t)^m(Q^m) = 1$, we have the equality if $\Gamma = Q^m$ in the above, and therefore

$$\sup_n \sum_{k>l} F_{n,k}(t, \Gamma) \leq \sup_n \sum_{k>l} F_{n,k}(t, Q^m) \to 0, \ l \to \infty, \ \Gamma \subset Q^m .$$

Thus we have by Proposition 2

$$\varliminf_{n\to\infty} v_n(t, \Gamma) \geq \varliminf_{n\to\infty} \sum_{k=1}^{n} F_{n,k}(t, \Gamma) = u(t)^m(\Gamma), \ \Gamma \subset Q^m ,$$

where $v_n(t, \Gamma)$ denotes the left hand side of (5.1) with $\varphi = \chi_\Gamma(x_1, \cdots, x_m)$. On the other hand,

$$\begin{aligned} \varlimsup_{n\to\infty} v_n(t, \Gamma) &= \varlimsup_{n\to\infty} \{v_n(t, Q^m) - v_n(t, Q^m - \Gamma)\} \\ &= 1 - \varliminf_{n\to\infty} v_n(t, Q^m - \Gamma) \leq 1 - u(t)^m(Q^m - \Gamma) = u(t)^m(\Gamma) , \end{aligned}$$

and so we must have $\lim_{n\to\infty} v_n(t, \Gamma) = u(t)^m(\Gamma)$, which was to be proved.

References

[1] Kac, M., Foundation of kinetic theory, Proc. Third Berkely Symp. on Math. Stat. and Prob. **3**, 171-197.

[2] McKean, H. P. Jr., A class of Markov processes associated with nonlinear parabolic equations, Proc. Nat. Acad. Sci. **56** (1966), 1907-1911.

[3] McKean, H. P. Jr., An exponential formula for solving Boltzmann's equation for a Maxwellian gas, J. Combinatorial Theory **2** (1967), 358-382.

[4] Johnson, D.P., On a class of stochastic processes and its relationship to infinite particle gases, Trans. Amer. Math. Soc. **132** (1968), 275-295.
[5] Tanaka, H., Propagation of chaos for certain Markov processes of jump type with nonlinear generators, I, II. Proc. Japan Acad. **45** (1969), 449-452, 598-600.
[6] Tanaka, H., Purely discontinuous Markov processes with nonlinear generators and their propagation of chaos, to appear in Теория Вероят. и ее Примен.
[7] Tanaka, S., An extension of Wild's sum for solving certain non-linear equation of measures, Proc. Japan Acad. **44** (1968), 884-889.
[8] Ueno, T., A class of Markov processes with bounded, non-linear generators, Japan. J. Math. **38** (1969), 19-38.
[9] Ueno, T., A class of purely discontinuous Markov processes with interactions, I, II. Proc. Japan Acad. **45** (1969), 348-353, 437-440.

(Received December 11, 1969)

Department of Mathematics
Faculty of Science
University of Tokyo
Hongo, Tokyo
113 Japan

Z. Wahrscheinlichkeitstheorie verw. Geb. 27, 47–52 (1973)

An Inequality for a Functional of Probability Distributions and Its Application to Kac's One-Dimensional Model of a Maxwellian Gas

Hiroshi Tanaka

1. Introduction

Let $\mathscr{P}$ be the class of 1-dimensional probability distributions f with

$$0<\alpha_2(f)<\infty,$$

where $\alpha_2(f)$ denotes the second moment of f. Taking a probability space (Ω, P) which is big enough to carry Gaussian random variables, we introduce a functional $\mathfrak{e}$ defined for $f\in\mathscr{P}$ by

$$\mathfrak{e}[f]=\inf E\{|X-Y|^2\},$$

where the infimum is taken over all pairs of random variables X and Y defined on (Ω, P) and distributed according to f and g respectively; here g is the Gaussian distribution with mean 0 and variance $\sigma^2=\alpha_2(f)$. $\mathfrak{e}[f]$ is sometimes denoted by $\mathfrak{e}[X]$ when X is a random variable with distribution f. It should be noticed that the value of $\mathfrak{e}[f]$ does not depend upon a choice of the probability space (Ω, P). The purpose of this paper is to present some basic properties of $\mathfrak{e}$ (especially, the inequality (2.2)) together with an application to the central limit theorem and then to show that the functional $\mathfrak{e}$ is monotone decreasing along Boltzmann solutions of Kac's one-dimensional model of a Maxwellian gas. Some of our results can be generalized to the case of R^3; for example, the functional $\mathfrak{e}$ similarly defined in R^3 decreases along solutions of Boltzmann's problem for the 3-dimensional Maxwellian gas, but this will be discussed in another occasion.

2. Basic Properties of $\mathfrak{e}$ and a Proof of the Central Limit Theorem

Theorem 1. *Let $f\in\mathscr{P}$, and denote by g the Gaussian distribution with mean* 0 *and variance $\sigma^2=\alpha_2(f)$. Let X and Y be random variables with distributions f and g, respectively. Then, $\mathfrak{e}[f]=E\{|X-Y|^2\}$ if and only if $X=F^{-1}(G(Y))$ almost surely, where F^{-1} is the right continuous inverse function of the distribution function F corresponding to f, and G is the distribution function corresponding to g.*

In the proof of this theorem the probability space is chosen as follows: Ω is the unit interval $[0, 1)$ and P is the Lebesgue measure in Ω. The proof is carried out in 3 steps.

Step 1. $(X(\omega)-X(\omega'))\,(Y(\omega)-Y(\omega'))\geqq 0$ for almost all (ω, ω') $(P\otimes P)$. Suppose the contrary holds. Then for some $\varepsilon>0$ at least one of the following events has positive $P\otimes P$-probability:

$$\{(\omega,\omega')\in\Omega\times\Omega:\ X(\omega)-X(\omega')<-3\varepsilon \text{ and } Y(\omega)-Y(\omega')>3\varepsilon\} \tag{1}$$

$$\{(\omega,\omega')\in\Omega\times\Omega:\ X(\omega)-X(\omega')>3\varepsilon \text{ and } Y(\omega)-Y(\omega')<-3\varepsilon\}. \tag{2}$$

We assume that the event (1) has positive probability for simplicity. Then, for some integers j_1, j_2, k_1, k_2 with $j_1+1<j_2$ and $k_1+1<k_2$, the event

$$\tilde{A}=\left\{(\omega,\omega'):\ \begin{array}{l} j_1\varepsilon<X(\omega)\leqq(j_1+1)\varepsilon,\ j_2\varepsilon<X(\omega')\leqq(j_2+1)\varepsilon \\ k_1\varepsilon<Y(\omega')\leqq(k_1+1)\varepsilon,\ k_2\varepsilon<Y(\omega)\leqq(k_2+1)\varepsilon \end{array}\right\}$$

has positive $P\otimes P$-probability. If we set

$$A=\{\omega:\ j_1\varepsilon<X(\omega)\leqq(j_1+1)\varepsilon,\ k_2\varepsilon<Y(\omega)\leqq(k_2+1)\varepsilon\}$$

$$A'=\{\omega:\ j_2\varepsilon<X(\omega)\leqq(j_2+1)\varepsilon,\ k_1\varepsilon<Y(\omega)\leqq(k_1+1)\varepsilon\},$$

then $\tilde{A}=A\times A'$ and hence $P(A)>0$, $P(A')>0$. Next, we take an irrational number λ and denote by T the Weyl automorphism: $\omega\in\Omega\to\omega+\lambda \pmod 1$. Then there exists an integer $n\geqq 0$ such that $P(A\cap T^{-n}A')>0$. If we set $B=A\cap T^{-n}A'$, $B'=T^nB$, $\varphi=T^n$, then $P(B)=P(B')>0$ and $B\cap B'=\emptyset$. We now define

$$X^{\#}(\omega)=\begin{cases} X(\varphi(\omega)) & \text{for } \omega\in B \\ X(\varphi^{-1}(\omega)) & \text{for } \omega\in B' \\ X(\omega) & \text{for } \omega\notin B\cup B'. \end{cases}$$

Since φ: $B\to B'$ is measure-preserving[1], $X^{\#}$ has still distribution f, and we have

$$\begin{aligned} E\{|X^{\#}-Y|^2\} &= \int_B |X(\varphi(\omega))-Y(\omega)|^2 P(d\omega)+\int_{B'} |X(\varphi^{-1}(\omega))-Y(\omega)|^2 P(d\omega) \\ &\quad + \int_{(B\cup B')^c} |X(\omega)-Y(\omega)|^2 P(d\omega) \\ &= \int_B \{|X(\varphi(\omega))-Y(\omega)|^2+|X(\omega)-Y(\varphi(\omega))|^2\} P(d\omega) \\ &\quad + \int_{(B\cup B')^c} |X(\omega)-Y(\omega)|^2 P(d\omega) \\ &< \int_B \{|X(\omega)-Y(\omega)|^2+|X(\varphi(\omega))-Y(\varphi(\omega))|^2\} P(d\omega) \\ &\quad + \int_{(B\cup B')^c} |X(\omega)-Y(\omega)|^2 P(d\omega)=E\{|X(\omega)-Y(\omega)|^2\}; \end{aligned}$$

the inequality part in the above employs the following elementary fact: if $a_1<a_2$ and $b_1<b_2$, then $(a_1-b_1)^2+(a_2-b_2)^2<(a_1-b_2)^2+(a_2-b_1)^2$. We thus arrive at a contradiction.

Step 2. Let $P(y,\cdot)$ be a regular conditional probability distribution of X given $Y=y$, and denote by S_y the smallest closed interval such that $P(y,S_y)=1$. We claim that

S_y and $S_{y'}$ are non-overlapping for almost all (y,y') with respect to $g\otimes g$. (2.1)

[1] I owe the use of the Weyl automorphism for constructing φ to Y. Takahashi.

Since $P(y,\cdot)\otimes P(y',\cdot)$ is a regular conditional probability distribution of $(X(\omega), X(\omega'))$ given $(Y(\omega), Y(\omega'))=(y,y')$, we have from Step 1

$$\iint g(dy)\,g(dy')\iint \chi(x,x',y,y')\,P(y,dx)\,P(y',dx')$$
$$=E\{\chi(X(\omega),X(\omega'),Y(\omega),Y(\omega'))\}=1$$

where χ is the indicator function of the set

$$\Gamma=\{(x,x',y,y')\in R^4:\ (x-x')(y-y')\geqq 0\}.$$

Therefore, for almost all (y,y') with respect to $g\otimes g$, we have

$$\iint \chi(x,x',y,y')\,P(y,dx)\,P(y',dx')=1.$$

So, if we set $\Gamma_1=\{(x,x')\in R^2: x\geqq x'\}$, then $P(y,\cdot)\otimes P(y',\cdot)$ is supported by Γ_1 for almost all $(y,y')\in\Gamma_1$; but this is a complicated way of saying that (2.1) holds.

Step 3. From Step 2 one can prove easily that S_y is a single point for almost all y with respect to g. Now, this fact combined with the inequality of Step 1 implies that X is an increasing function of Y (a.s.); this is possible only when $X=F^{-1}(G(Y))$ almost surely. The "if" part is obvious, since the infimum in the definition of $\mathfrak{e}[f]$ is actually attained by some pair.

Theorem 2. *Let X and Y be independent random variables with distributions f_1 and $f_2\in\mathscr{P}$, respectively, and assume that $E\{X\}=E\{Y\}=0$. Then, for any real constants a, b such that $a\neq 0$, $b\neq 0$,*

$$\mathfrak{e}[aX+bY]<a^2\,\mathfrak{e}[X]+b^2\,\mathfrak{e}[Y], \tag{2.2}$$

unless both X and Y are Gaussian.

The proof of this theorem is based upon Theorem 1. It is obvious that

$$\mathfrak{e}[aX+bY]\leqq a^2\,\mathfrak{e}[X]+b^2\,\mathfrak{e}[Y]$$

holds, and so assuming the equality holds in the above, we will prove $f_i=g_i$, where g_i is the Gaussian distribution with mean 0 and variance $\sigma_i^2=\alpha_2(f_i)$, $i=1,2$. If X_1 and X_2 are independent random variables with distributions g_1 and g_2, respectively, then with the obvious notation it follows from Theorem 1 that

$$a^2\,\mathfrak{e}[X]+b^2\,\mathfrak{e}[Y]=a^2\,E\{|F_1^{-1}(G_1(X_1))-X_1|^2\}+b^2\,E\{|F_2^{-1}(G_2(X_2))-X_2|^2\}$$
$$=E\{|aF_1^{-1}(G_1(X_1))+bF_2^{-1}(G_2(X_2))-(aX_1+bX_2)|^2\}.$$

Since aX_1+bX_2 is also G-distributed, we have again from Theorem 1

$$aF_1^{-1}(G_1(X_1))+bF_2^{-1}(G_2(X_2))=F^{-1}(G(aX_1+bX_2))\ \text{a.s.}, \tag{2.3}$$

where F is the distribution function of $aX+bY$. By the right continuity of the functions involved, (2.3) yields

$$aF_1^{-1}(G_1(x))+bF_2^{-1}(G_2(y))=F^{-1}(G(ax+by))$$

for all $x,y\in R^1$. This functional equation for unknown F_1, F_2, F can easily be solved; the result is $F_1=G_1$, $F_2=G_2$, completing the proof.

We next list some simple properties of $\mathfrak{e}$ for later use.

1. If f_n converges to some $f \in \mathscr{P}$ as $n \uparrow \infty$ in such a way that

$$\lim_{N\to\infty} \sup_{n\geq 1} \int_{|x|>N} x^2 f_n(dx) = 0, \tag{2.4}$$

then $\lim_{n\to\infty} \mathfrak{e}[f_n] = \mathfrak{e}[f]$. The condition (2.4) is satisfied if for some $p>2$ the absolute p-th moments of f_n are bounded in n.

2. Let $f_\theta \in \mathscr{P}$ and $\alpha_2(f_\theta) = \sigma^2$ for $0 \leqq \theta < 1$, and assume that $\int \varphi(x) f_\theta(dx)$ is Borel measureable in θ for any bounded continuous function φ. Then for any probability measure μ on $[0, 1)$ we have

$$\mathfrak{e}\left[\int f_\theta\, \mu(d\theta)\right] \leqq \int \mathfrak{e}[f_\theta]\, \mu(d\theta).$$

3. By Theorem 1, $\mathfrak{e}$ admits the expression

$$\mathfrak{e}[f] = 2 \int \{x^2 - x\, G^{-1}(F(x))\}\, f(dx)$$

for a continuous probability distribution f in $\mathscr{P}$.

The inequality (2.2) will now be applied to give a simple proof of the central limit theorem. Let $\{X_n\}_{n=1,2,\dots}$ be a sequence of independent identically distributed random variables with mean 0 and variance 1. Then the so-called central limit theorem states that the distribution of $\xi_n = n^{-\frac{1}{2}}(X_1 + \dots + X_n)$ tends to a Gaussian distribution as $n \uparrow \infty$. *Here we prove that* $\mathfrak{e}[\xi_n] \to 0$ *as* $n \to \infty$ *assuming* $E\{X_1^4\} < \infty$[2]. This condition implies that

$$E\{\xi_n^4\} = \frac{1}{n} E\{X_1^4\} + 3\left(1 - \frac{1}{n}\right) < \text{const.} \quad \text{(independent of } n\text{)}. \tag{2.5}$$

Putting $\eta_k = \xi_{2^k}$, we first prove that $\mathfrak{e}[\eta_k] \downarrow 0$ as $k \uparrow \infty$. The decreasing property of $\mathfrak{e}[\eta_k]$ is obvious by the inequality (2.2), and so we denote by l the limit of $\mathfrak{e}[\eta_k]$ as $k \to \infty$. If f is a limit distribution of η_k as $k \to \infty$ via some subsequence $k_1 < k_2 < \cdots$, and if η and ζ are independent random variables with distribution f, then it follows from (2.5) and 1 of § 2 that

$$\mathfrak{e}[\eta] = \lim_{p\to\infty} \mathfrak{e}[\eta_{k_p}] = l \quad \text{and} \quad \mathfrak{e}\left[\frac{\eta+\zeta}{2}\right] = \lim_{p\to\infty} \mathfrak{e}[\eta_{2^{k_p}}] = l,$$

therefore by Theorem 2 the limit l must be 0. Next, we write an integer $n \geqq 1$ as $n = \sum_{k=0}^{m} n_k$ where $n_k = \varepsilon_k 2^k$ with $\varepsilon_k = 0$ or 1. Then, using the inequality (2.2) we have $\mathfrak{e}[\xi_n] \leqq \frac{1}{n} \sum_{k=0}^{m} n_k\, e[\eta_k]$, and hence $\mathfrak{e}[\xi_n] \to 0$ as was to be proved.

3. $\mathfrak{e}$ Decreases along Solutions of Boltzmann's Problem for Kac's Model of a Maxwellian Gas

Given $f_1, f_2 \in \mathscr{P}$ and $\theta \in [0, 2\pi)$, we denote by $B_\theta(f_1, f_2)$ the probability distribution of $X_1 \cos\theta + X_2 \sin\theta$, where X_1 and X_2 are random variables with distri-

[2] This condition is assumed just to simplify the proof. Without this $\mathfrak{e}[\xi_n]$ still tends to 0.

butions f_1 and f_2 respectively. We also put

$$B(f_1,f_2)=\frac{1}{2\pi}\int_0^{2\pi} B_\theta(f_1,f_2)\,d\theta.$$

In Kac's one-dimensional model of a Maxwellian gas, the distribution $u(t,dx)$ of molecular speeds at time $t>0$ is determined by the solution of Boltzmann's problem

$$\frac{\partial u(t,\cdot)}{\partial t}=B\big(u(t),u(t)\big)-u(t,\cdot). \tag{3.1}$$

The solutions of this equation can be obtained by Wild's sum [2]. If the initial distribution has a density, then so does the solution, and it is known that the entropy increases along the solution with time, while the solution itself tends to a Gaussian distribution as $t\uparrow\infty$. McKean [1] gave detailed discussions on this subject; he gave also other functionals which are (or, at least are expected to be) monotone along the solutions of (3.1) together with an interesting conjecture about them. But, among these functionals, the entropy and Linnik's functional are the only ones which were used effectively in the investigation of the asymptotic properties of the solutions of (3.1). In this section, we prove that the functional e decreases monotonically to zero along the solutions of (3.1); this statement itself implies automatically that the solutions of (3.1) tend to Gaussian distributions as $t\uparrow\infty$.

Theorem 3. *Let $u(t)$ be the solution of* (3.1) *with initial distribution $f\in\mathscr{P}$. Then,* (i) $\mathrm{e}[u(t)]$ *is decreasing in t, and* (ii) *if f has finite fourth moment,* $\mathrm{e}[u(t)]$ *decreases to* 0 *as $t\uparrow\infty$.*

The following corollary is an immediate consequence of the above theorem and 3 of § 2.

Corollary. *Let $\mathscr{P}_0$ be the subclass of $\mathscr{P}$ consisting of continuous probability distributions, and put $\mathrm{e}_0[f]=\int x\,G^{-1}[F(x)]\,f(dx)$. Then the functional e_0 is increasing along the solutions of* (3.1) *with initial distributions $\in\mathscr{P}_0$.*

The proof of Theorem 3 will be given in several steps.

Proof. 1. Let $\mathscr{P}_n(f)$, $n\geqq 1$, be the (finite) set of probability measures from $\mathscr{P}$ defined inductively as follows: (i) $\mathscr{P}_1(f)$ consists of a single element f, and (ii) $\mathscr{P}_n(f)$ is the set of all probability measures of the form $B(f_1,f_2)$ with $f_1\in\mathscr{P}_{n_1}(f)$, $f_2\in\mathscr{P}_{n_2}(f)$, $n_1+n_2=n$. Then, the solution $u(t)$ of (3.1) with initial distribution f can be expressed as Wild's sum

$$u(t)=\mathrm{e}^{-t}\sum_{n=1}^{\infty}(1-\mathrm{e}^{-t})^{n-1}\,p_n(f), \tag{3.2}$$

where $p_n(f)$ stands for a convex combination of elements in $\mathscr{P}_n(f)$, $n\geqq 1$ ([2], see also [1]).

2. If $\tilde f$ denotes the even part of f, say $\tilde f(dx)=\frac{1}{2}\big(f(dx)+f(-dx)\big)$, then it is easy to see that $B(f_1,f_2)=B(\tilde f_1,\tilde f_2)$. Therefore, if f_1 and f_2 have the same second

moment, it follows from 2 of § 2 and Theorem 2 that

$$\mathrm{e}[B(f_1, f_2)] = \mathrm{e}[B(\tilde{f}_1, \tilde{f}_2)] \leqq \frac{1}{2\pi} \int_0^{2\pi} \mathrm{e}[B_\theta(\tilde{f}_1, \tilde{f}_2)]\, d\theta$$
$$\leqq \frac{1}{2\pi} \int_0^{2\pi} \{\mathrm{e}[\tilde{f}_1] \cos^2\theta + \mathrm{e}[\tilde{f}_2] \sin^2\theta\}\, d\theta \leqq \frac{\mathrm{e}[f_1] + \mathrm{e}[f_2]}{2},$$

because $\mathrm{e}[\tilde{f}] \leqq \mathrm{e}[f]$. Therefore we have $\mathrm{e}[p_n(f)] \leqq \mathrm{e}[f]$, and hence by Wild's sum (3.2) and 2 of § 2 we see that

$$\mathrm{e}[u(t)] \leqq \mathrm{e}[f] \qquad (t > 0); \tag{3.3}$$

the equality holds if and only if f is a Gaussian distribution. (3.3) implies the part (i) of the theorem.

3. If $\int x^4 f(dx) < \infty$, then by (3.1) the function $\alpha(t) = \int x^4 u(t, dx)$ satisfies the differential equation

$$\frac{d\alpha(t)}{dt} = \frac{3}{4}\sigma^4 - \frac{1}{4}\alpha(t), \qquad \sigma^2 = \alpha_2(f),$$

which implies that $\alpha(t) \to 3\sigma^4$ as $t \to \infty$, and hence $\alpha(t)$ is bounded. Next, let u_∞ be a limit distribution of $u(t)$ as $t \uparrow \infty$ via some subsequence $t_1 < t_2 < \cdots$. Since $\alpha(t)$ is bounded, we have $\mathrm{e}[u_\infty] = \lim_{n\to\infty} \mathrm{e}[u(t_n)] = \lim_{t\to\infty} \mathrm{e}[u(t)]$ by 1 of § 2. If $u_\infty(t)$ denotes the solution of (3.1) with initial distribution u_∞, then an application of Wild's sum shows that $u_\infty(t) = \lim_{n\to\infty} u(t_n + t)$ and hence $\mathrm{e}[u_\infty(t)] = \lim_{t\to\infty} \mathrm{e}[u(t_n + t)] = \mathrm{e}[u_\infty]$. Therefore u_∞ must be a Gaussian distribution from the preceding step, as was to be proved.

Note. After sending the manuscript to the editor, I was informed from T. Yanagimoto of a simple proof of Theorem 1 based upon the following Hoeffding's formula: if F denotes the joint and F_X and F_Y the marginal distribution functions of X and Y, then

$$E(XY) - E(X)\,E(Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [F(x, y) - F_X(x)\, F_Y(y)]\, dx\, dy$$

provided the expectations on the left hand side exist.

References

1. McKean, H. P.: Speed of approach to equilibrium for Kac's caricature of a Maxwellian gas. Arch. Rat. Mech. Analysis **21**, 343–367 (1966)
2. Wild, E.: On Boltzmann's equation in the kinetic theory of gases. Proc. Cambridge Philos. Soc. **47**, 602–609 (1951)

Hiroshi Tanaka
Department of Mathematics
Hiroshima University
Hiroshima, Japan

(Received January 10, 1973)

ON MARKOV PROCESS CORRESPONDING TO BOLTZMANN'S EQUATION OF MAXWELLIAN GAS

Hiroshi Tanaka

§1. Introduction. The basic equation in the kinetic theory of dilute monoatomic gases is the famous Boltzmann's equation. In the spatially homogeneous case, the initial value problem of this equation was solved for a gas of hard balls by Carleman [1], for Maxwellian gas with cutoff by Wild [14], and for bounded total collision cross-section by Povzner [8] (in modified spatially inhomogeneous case), but it seems that no results (for existence and uniqueness) have been obtained for Maxwellian gas without cutoff. On the other hand, H. P. McKean [5] introduced a class of Markov processes associated with certain nonlinear (parabolic) equations such as Boltzmann's equation, and brought a new light in the field of investigation of such equations by probabilistic methods (see also [6]). Then, there appeared works by D. P. Johnson [3], T. Ueno [11] [12], Y. Takahashi [9] and H. Tanaka [10], mostly concerned with Boltzmann's equation of cutoff type and certain nonlinear equations with similar structure. Especially, Ueno [12] constructed Markov processes which describe motions of infinitely many interacting particles, while Takahashi [9] introduced interaction semigroups and discussed their relationship to branching semigroups. In this paper we are exclusively concerned with non-cutoff Maxwellian gas ; our purpose is to construct a Markov process in the sense of McKean [5] corresponding to the 3-dimensional Maxwellian gas without cutoff by solving appropriate stochastic differential equation (the equation (2.10) in §2). The theory of stochastic differential equations was initiated by K. Itô [2] and, in the case of diffusions, equations similar to (2.10) were considered by McKean [7] in connection with certain nonlinear parabolic equations. The results are only summarized ; full proofs will be published elsewhere.

We consider a monoatomic dilute gas composed of a large number of molecules moving in the space and assume that there are no outside forces. Let $Nu(t,x)dx$ be the number of molecules with velocities x within the differential element dx at time t, where N is the total number of molecules. Then under the assumption of spatial homogeneity, $u(t,x)$ satisfies the following Boltzmann's equation :

$$(1.1) \quad \frac{\partial u(t,x)}{\partial t} = \int_{S_o \times R^3} \{u(t,x^*)u(t,y^*) - u(t,x)u(t,y)\}|x-y|Q(|x-y|,\theta)\sin\theta d\theta d\psi dy,$$

where $S_o = (0,\pi)\times[0,2\pi)$ and θ,ψ are points in $(0,\pi)$ and $[0,2\pi)$ respectively. Denote by $S_{x,y}$ the sphere with center $\frac{x+y}{2}$ and diameter $|x-y|$, and on this sphere we consider a spherical coordinate system with polar axis defined by the relative velocity $x-y$. x^* and y^* are the post-collisional velocities. Let θ and ψ be the colatitude of x^*(the angle between two vectors $x-y$ and x^*-y^*) and the logitude of x^*, respectively. By the conservation laws of momentum and energy x^* and y^* are always situated on $S_{x,y}$ and constitute a diameter of $S_{x,y}$, and so y^* is also determined by θ and ψ. For each x and y the origin of the longitude ψ may be arbitrary chosen within the requirement that x^* and y^* as functions of (x,y,θ,ψ) should be Borel measurable. A nonnegative function Q is determined by the intermolecular force and is called the differential collision cross-section. In the model of gas of hard balls Q is a positive constant, while in the Maxwellian model in which molecules repel each other with a force inversely proportional to the fifth power of their distance, $|x-y|Q(|x-y|,\theta)$ turns out to be a function $Q_M(\theta)$ of θ alone ; in the latter case $Q_M(\theta)$ is a decreasing function of θ with $Q_M(\theta) \sim \text{const.}\ \theta^{-\frac{5}{2}}$, $\theta\downarrow 0$, and so the total collision cross-section is infinite (non cutoff) (see [13]). This is the case we consider in this paper.

§2. Markov processes and stochastic differential equation

In order to indicate our problem clearly, we first explain how a Markov process in the sense of McKean [5] is associated with Boltzmann's equation, taking gas of hard balls by example. The equation (1.1) for gas of hard balls is usually treated in the following form :

$$\frac{\partial u(t,x)}{\partial t} = \int_{S^2\times R^3} \{u(t,x^*)u(t,y^*)-u(t,x)u(t,y)\}|(y-x),\ell)|d\ell dy, \tag{2.1}$$

where $x^*=x+(y-x,\ell)\ell$, $y^*=y-(y-x,\ell)\ell$, $\ell\in S^2$, and $d\ell$ is the uniform distribution on S^2. We set

$$\begin{cases} u(t,\Gamma) = \int_\Gamma u(t,x)dx \quad , \quad \Gamma\in\mathcal{B}(R^3) \\ u(t,\varphi) = \int \varphi(x)u(t,dx) \quad , \quad \varphi\in C_b(R^3) \end{cases} \tag{2.2}$$

where $C_b(R^3)$ denotes the space of real valued bounded continuous functions on R^3 (the notation (2.2) will be used throughout in this paper). Then, from (2.1) we have

$$\frac{\partial u(t,\varphi)}{\partial t} = \int_{S^2\times R^3\times R^3} \{\varphi(x^*)-\varphi(x)\}|(y-x,\ell)|d\ell u(t,dx)u(t,dy), \ \varphi\in C_b(R^3). \tag{2.3}$$

Povzner's result [8] may be stated as follows : given an initial data f (probability measure on R^3) such that $\int|x|^4f(dx)<\infty$, there exists a unique solution $u(t,\cdot)$ (probability measure) of (2.3) such that $\mu(t)=\int|x|^4u(t,dx)$ is locally bounded. Now, keeping f and $u(t,\cdot)$ as above, we consider the following equation for $v(t,\cdot)$:

$$(2.4)\quad \begin{cases} \dfrac{\partial v(t,\varphi)}{\partial t} = \displaystyle\int_{S^2\times R^3\times R^3} \{\varphi(x^*)-\varphi(x)\}\,|(y-x,\ell)|\,d\ell\, v(t,dx)\,u(t,dy) \\ v(0,\cdot) = \delta(z,\cdot), \qquad \varphi\in C_b(R^3) , \end{cases}$$

where $\delta(z,\cdot)$ denotes the unit distribution with mass at z. Then, using the result of [8] it is not hard to prove that (2.4) has a unique solution, which we denote by $P_f(t,z,\cdot)$. We can also prove that

(2.5a) $\qquad P_f(t,z,\cdot)$ is a probability measure on R^3,

$$(2.5b)\qquad u(t,\Gamma) = \int_{R^3} P_f(t,z,\Gamma)f(dz), \quad \Gamma\in\mathcal{B}(R^3) ,$$

$$(2.5c)\qquad P_f(t+s,z,\Gamma) = \int_{R^3} P_f(t,z,dy)P_{u(t)}(s,y,\Gamma).$$

The above properties of $P_f(t,z,\cdot)$ enable us to construct a Markov process as follows. Let Ω be the space of step functions $z(t)$ defined on $[0,\infty)$ and taking values in R^3, and $\mathcal{B}$ the σ-field generated by (measurable) cylinder sets in Ω. For each probability measure f on R^3 such that $\int|x|^4 f(dx)<\infty$, we can construct a probability measure $P_f(\cdot)$ on $(\Omega,\mathcal{B})$ so that $\{\Omega,z(t),P_f\}$ is a Markov process in the following sense.

$$(2.6a)\qquad P_f\{z(0)\in dx\} = f(dx) ,$$

$$(2.6b)\qquad P_f\{z(t+s)\in\Gamma\,|\,z(\tau):0\leqq\tau\leqq t\}=P_{u(t)}(s,z(t),\Gamma), \text{ a.s. } (P_f).$$

This is the Markov process in the sense of [5] associated with the gas of hard balls (2.1).

Now, our problem can be stated as follows : construct a Markov process $x(t)$ in the sense of [5] which is related to the Maxwellian gas without cutoff as $z(t)$ is to (2.1). So we take up the equation (1.1) with Q_M specified as in §1 and rewrite it as in the form (2.3). A formal but careful calculation yields

$$\frac{\partial u(t,\varphi)}{\partial t} = \int_{S_o \times R^3 \times R^3} \{\varphi(x^*) - \varphi(x))Q_M(\theta)\sin\theta d\theta d\psi u(t,dx)u(t,dy). \tag{2.7}$$

Considering the singularity of $Q_M(\theta)$ at $\theta=0$, it may be natural to take φ (test function) from the space $C_o^1(R^3)$ of C^1-functions with compact supports. However, we do not treat (2.7) directly ; instead we consider a suitable stochastic differential equation as will be described below.

Let $S=(0,1)\times(0,\pi)\times[0,2\pi)$ and λ the measure on S defined by $d\lambda = d\alpha\otimes Q(d\theta)\otimes d\psi$, where $Q(d\theta)=Q_M(\theta)\sin\theta d\theta$. Although we have thus specified $Q(d\theta)$, we remark that the only property of $Q(d\theta)$ we need later is

$$\int_{(0,\pi)} \theta Q(d\theta) < \infty, \tag{2.8}$$

and so our results remain valid for arbitrary measure $Q(d\theta)$ subject to the condition (2.8). Let $\tilde{\lambda}$ be the product measure $dt\otimes d\lambda$ on the space $(0,\infty)\times S$ and denote by $\mathcal{F}$ the class of Borel sets in $(0,\infty)\times S$ which have finite $\tilde{\lambda}$-measure. A family of random variables $\{p(A,\omega)\}_{A\in\mathcal{F}}$ defined on a probability space $(\Omega,\mathcal{B},P)$ is called a Poisson random measure on $(0,\infty)\times S$ associated with the measure $\tilde{\lambda}$, if the following two conditions are satisfied.

(i) Each $p(A,\omega)$ is distributed according to the Poisson distribution with mean $\tilde{\lambda}(A)$.

(ii) If $A_1, A_2, \cdots \in \mathcal{F}$, $A_j \cap A_K = \phi\ (j \neq k)$ and $A = \bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$,

then the family $\{p(A_k,\omega)\}_{k=1,2,\cdots}$ is independent and

$$p(A,\omega) = \sum_{k=1}^{\infty} p(A_k,\omega) \quad \text{(a.s.)}. \tag{2.9}$$

Since a Poisson random measure always admits a suitable modification for which (2.9) holds for <u>all</u> ω, we may suppose, if necessary, that $\{p(A,\omega)\}$ satisfies this stronger version of (2.9). For a Poisson random measure $\{p(A,\omega)\}_{A\in\mathcal{F}}$ we set

$$\mathcal{B}^t_\infty = \sigma\{p(A,\omega):\ A \subset (t,\infty)\times S\},$$

where the notation $\sigma\{$ —— $\}$ stands for the smallest σ-field that makes all the random variables in $\{$ —— $\}$ measurable. Next, we regard the unit interval (0,1) as a probability space by considering the Lebesgue measure (precisely, its restriction) on the Borel field of (0,1), and on this probability space we sometimes consider an R^3-valued stochastic process $\{y(t,\alpha),\ 0 \leq t < \infty\}$ with path functions which are right continuous and have left hand limits. Such a process is called an <u>α-process</u> for simplicity ; similarly a random variable defined on the probability space (0,1) is called an α-random variable. Now our stochastic differential equation associated with (2.7) can be written as follows :

$$x(t,\omega) = x(0,\omega) + \int_{(0,t]\times S} a(x(s,\omega), y(s,\alpha), \theta, \psi)\, p(ds d\sigma,\omega). \tag{2.10a}$$

(2.10b) $\{y(t,\alpha),\ 0 \leq t < \infty\}$ is an α-process which is equivalent in law to $\{x(t,\omega),\ 0 \leq t < \infty\}$.

(2.10c) For each $t \geq 0$, the σ-field

$$\sigma\{x(s,\omega),\ p(A,\omega)\ :\ 0 \leq s \leq t,\ \ A \subset (0,t]\times S\}$$

is independent of $\mathcal{B}^t_\infty$

Here $a(x,y,\theta,\psi)=x^*(x,y,\theta,\psi)-x$, $d\sigma=d\alpha d\theta d\psi$, and of course $\{p(A,\omega)\}$ is a Poisson random measure associated with $\tilde{\lambda}$ defined over a (basic) probability space $(\Omega,\mathcal{B},P)$. $\{x(t,\omega),\ 0\leq t\leq\infty\}$ is unknown process to be determined by the equation (2.10a) under the additional conditions (2.10b) and (2.10c). Solving (2.10) gives rise to an answer to the problem of finding Markov process associated with the Maxwellian gas, as will be seen in the next section.

§3. Solving the stochastic differential equation. Main results

Let f be a probability measure on R^3. By a solution of (2.10) with initial distribution f, we mean an R^3-valued stochastic process $\{x(t,\omega),\ 0\leq t\leq\infty\}$ defined over a suitable probability space $(\Omega,\mathcal{B},P)$ and satisfying the following conditions:

(i) $x(t,\omega)$ is right continuous and has left hand limits with probability 1,

(ii) $P\{x(0,\omega)\in dx\} = f(dx)$,

(iii) the relations (2.10a), (2.10b) and (2.10c) hold for some Poisson random measure $\{p(A,\omega)\}$ associated with $\tilde{\lambda}$ and defined over the probability space $(\Omega,\mathcal{B},P)$.

In solving the existence problem, the usual method of successive approximation can not be applied directly, since the function $a(x,y,\theta,\psi)$ is not smooth. However, it can be shown that $a(x,y,\theta,\psi)$ has a nice property similar to the Lipschitz continuity (Lemma 1). Owing to this property a kind of successive approximation method works, but in the results the usual pathwise uniqueness will be replaced by weaker one, the uniqueness in law. Here, we say that the uniqueness in law holds for initial distribution f, if any two solutions of (2.10) with initial distribution f (which may be defined on different probability spaces) are identical in law as stochastic processes.

In what follows, $a(x,y,\theta,\psi)$ viewed as function of ψ alone is regarded as a periodic function on R^1 with period 2π.

<u>Lemma 1</u>. <u>There exists a Borel function</u> $\psi_o(x,y,\tilde{x},\tilde{y},)$ of $x,y,\tilde{x},\tilde{y} \in R^3$ <u>such that</u>

$$|a(x,y,\theta,\psi)-a(\tilde{x},\tilde{y},\theta,\psi+\psi_o(x,y,\tilde{x},\tilde{y}))| \leqq K\{|x-\tilde{x}|+|y-\tilde{y}|\}\cdot\theta,$$

<u>where</u> K <u>is an absolute constant</u>.

In fact, ψ_o can be defined as follows. If $x=y$ or $\tilde{x}=\tilde{y}$, we set $\psi_o(x,y,\tilde{x},\tilde{y}) =0$. If $x\neq y$ and $\tilde{x}\neq\tilde{y}$, we set

$$x^o = \frac{|x-y|}{|\tilde{x}-\tilde{y}|}\ \frac{\tilde{x}-\tilde{y}}{2} + \frac{x+y}{2}$$

and denote by ρ the rotation around ℓ such that $\rho x = x^o$, where ℓ is the straight line passing through the point $\frac{x+y}{2}$ and perpendicular to the plane determined by the three points $\frac{x+y}{2}$, x, x^o (when $x^o=y$, we define ρ by $\rho z = x+y-z$). Now we set

$$x^{\#}(x,y,\theta,0)= \frac{|\tilde{x}-\tilde{y}|}{|x-y|}\ (\rho x^*(x,y,\theta,0)- \frac{x+y}{2}) + \frac{\tilde{x}+\tilde{y}}{2}\ .$$

Then $x^{\#}$ lies on the sphere $S_{\tilde{x},\tilde{y}}$ and hence we can define ψ_o $(0\leqq\psi_o<2\pi)$ by the formula

$$x^{\#}(x,y,\theta,0) = x^*(\tilde{x},\tilde{y},\theta,\psi_o(x,y,\tilde{x},\tilde{y})).$$

One of our main results is

<u>Theorem A</u>. <u>Assume that</u> $\int |x| f(dx)<\infty$. <u>Then</u>,

(i) <u>there exists a solution</u> $\{x(t,\omega),\ 0\leqq t<\infty\}$ <u>of</u> (2.10) <u>with initial distribution</u> f,

(ii) <u>the uniqueness in law holds for initial distribution</u> f,

(iii) <u>the probability distribution</u> $u(t,\cdot)$ <u>of</u> $x(t,\omega)$ <u>satisfies</u> (2.7) <u>for any</u> $\varphi \in C_o^1(R^3)$,

(iv) (a) (<u>conservation of momentum</u>) $E\{x(t,\omega)\}$ <u>is independent of</u> t,

(b) (<u>conservation of energy</u>) $E\{|x(t,\omega)|^2\}$ <u>is independent of</u> t, <u>provided</u> $\int |x|^2 f(dx)<\infty$.

Here is an outline of the proof of the existence part (i). Choose a sequence $\{\xi_n(\alpha)\}_{n=o,1,\cdots}$ of independent α-random variables with the uniform distribution on $(0,1)$, and set $\mathcal{F}_n = \sigma\{\xi_k(\alpha),\ 0\leqq k\leqq n\}$. Over a suitable probability space $(\Omega,\mathcal{B},P)$ we take an R^3-valued random variable $x(0,\omega)$ with distribution f and a Poisson random measure $\{\tilde{p}(A,\omega)\}$ associated with $\tilde{\lambda}$ so that $x(0,\omega)$ is independent of $\{\tilde{p}(A,\omega)\}$, and set $x_o(t,\omega)=x(0,\omega)$ for all $t\geqq 0$. Then, taking arbitrary $\mathcal{F}_o$--measurable α-process $\{y_o(t,\alpha),\ 0\leqq t<\infty\}$ which is equivalent in law to the process $\{x_o(t,\omega),\ 0\leqq t<\infty\}$, we put

$$x_1(t,\omega)=x(0,\omega)+\int_{(0,t]\times S} a(x_o(s,\omega),y_o(s,\alpha),\theta,\psi)\tilde{p}(dsd\sigma,\omega).$$

In general, for $n\geqq 1$ we put

$$x_{n+1}(t,\omega)=x(0,\omega)+\int_{(0,t]\times S} a(x_n(s,\omega),y_n(s,\omega),\theta,\psi+\tilde{\psi}_n)\tilde{p}(dsd\sigma,\omega),$$

where $\{y_n(t,\alpha),\ 0\leqq t<\infty\}$ is an $\mathcal{F}_n$-measurable *) α-process such that the process $\{(y_{n-1}(t,\alpha),y_n(t,\alpha)),\ 0\leqq t<\infty\}$ is equivalent in law to the process $\{(x_{n-1}(t,\omega),$ $x_n(t,\omega)),\ 0\leqq t<\infty\}$, and $\tilde{\psi}_n=\tilde{\psi}_{n-1}+\psi_o(x_{n-1}(s,\omega),y_{n-1}(s,\alpha),x_n(s,\omega),y_n(s,\alpha))$, $\tilde{\psi}_o=0$.

Then by Lemma 1

$$(3.1)\qquad |a(x_{n-1}(s),y_{n-1}(s),\theta,\psi+\tilde{\psi}_{n-1})-a(x_n(s),y_n(s),\theta,\psi+\tilde{\psi}_n)| \leqq K\{|x_n(s)-x_{n-1}(s)|+|y_n(s)-y_{n-1}(s)|\}\theta,$$

and hence

$$E|x_{n+1}(t)-x_n(t)|\leqq \text{const.}\int_o^t E|x_n(s)-x_{n-1}(s)|ds,$$

which implies

$$\sum_{n=1}^{\infty} E|x_n(t)-x_{n-1}(t)|<\infty.$$

*) *$\mathcal{F}_n$-measurability is imposed only to make the construction of $\{y_n(t,\alpha)\}$ easier.*

Therefore, $x(t)=\lim_{n\to\infty} x_n(t)$ exists (a.s.), and hence the same for $y(t)=\lim_{n\to\infty} y_n(t)$. Again using (3.1), it can be shown that $\lim_{n\to\infty} a(x_n(t),y_n(t),\theta,\psi+\tilde{\psi}_n)$ exists and is equal to $a(x(t),y(t),\theta,\psi+\tilde{\psi}_\infty)$ with some $\tilde{\psi}_\infty=\tilde{\psi}_\infty(t,\alpha,\omega)$ except on a negligible set. Now setting

$$p(A,\omega)=\int_{(0,\infty)\times S} \chi_A(t,\alpha,\theta,\psi+\tilde{\psi}_\infty)\tilde{p}(dtd\sigma,\omega),$$

we obtain (2.10).

The uniqueness part (ii) follows from the following

Lemma 2. Let f *be the same as in the theorem and* $\{x(t,\omega),\ 0\leq t<\infty\}$ *any solution of* (2.10) *with initial distribution* f. *Take arbitrary finite interval* $[0,T]$ *and let* Δ *be a partition of* $[0,T]$: $0=t_0<t_1<\cdots<t_n=T$. *Choosing an* R^3*-valued random variable* $X(0)$ *independent of* $\{p(A)\}$ *and with distribution* f, *we define a process* $\{X_\Delta(t),\ 0\leq t\leq T\}$ *as follows* :

$$\begin{cases} X_\Delta(t)=X(0)+\int_{(0,t]\times S} a(X(0),Y_0(\alpha),\theta,\psi)p(dsd\sigma), & 0\leq t\leq t_1 \\ X_\Delta(t)=X_\Delta(t_1)+\int_{(t_1,t)\times S} a(X_\Delta(t_1),Y_1(\alpha),\theta,\psi)p(dsd\sigma), & t_1<t\leq t_2 \\ \quad, & \\ \quad, & \\ \quad, & \\ X_\Delta(t)=X_\Delta(t_{n-1})+\int_{(t_{n-1},t]\times S} a(X_\Delta(t_{n-1}),Y_{n-1}(\alpha),\theta,\psi)p(dsd\sigma), & t_{n-1}<t\leq T, \end{cases}$$

where $Y_0(\alpha),\cdots,Y_{n-1}(\alpha)$ *are* α*-random variables defined successively so that each* $Y_k(\alpha)$ *has the same distribution as* $X_\Delta(t_k)$. *Then, any finite dimensional distribution of* $\{X_\Delta(t),\ 0\leq t\leq T\}$ *converges to the corresponding finite dimensional distribution of* $\{x(t),\ 0\leq t\leq T\}$ *as* $|\Delta|=\max_{1\leq k\leq n}|t_k-t_{k-1}|$ *tends to* 0.

The following theorem shows that the process $\{x(t,\omega),\ 0\leq t<\infty\}$ is the Markov process we were looking for. The proof is almost the same as that of Theorem A.

Theorem B. (i) *Let* $\{y(t,\alpha),\ 0\leq t<\infty\}$ *be any* α-*process equivalent in law to the process* $\{x(t,\omega),\ 0\leq t<\infty\}$ *constructed in Theorem A. Then, for each* $x\in R^3$ *there exists a solution* $z(t,\omega)$ *for the stochastic differential equation*

$$z(t,\omega)=x+\int_{(0,t]\times S} a(z(s,\omega),y(s,\alpha)\theta,\psi)\,p(ds\,d\sigma,\omega),$$

and the uniqueness in law holds for this equation.

(ii) *Set*

$$P_f(t,x,\Gamma)=P\{z(t,\omega)\in\Gamma\}\ ,\quad \Gamma\in\mathcal{B}(R^3)\ ,$$

and let $\{x(t,\omega),\ 0\leq t<\infty\}$ *be the solution of* (2.10). *Then*

$$P\{x(t+s,\omega)\in\Gamma\,|\,\mathcal{B}_t\}=P_{u(t)}(s,x(t,\omega),\Gamma)\quad \text{a.s.},$$

where $\mathcal{B}_t=\sigma\{x(s,\omega),p(A,\omega)\ :\ 0\leq s\leq t,\ \ A\subset(0,t]\times S\}$.

References

[1] T. Carleman, Problèmes Mathematiques dans la Théorie Cinétique des Gaz, Uppsala, 1957.

[2] K. Itô, On stochastic differential equations, Mem. Amer. Math. Soc. No. 4 (1951).

[3] D. P. Johnson, On a class of stochastic processes and its relationship to infinite particle gases, Trans. Amer. Math. Soc. 132(1968),275-295.

[4] M. Kac, Foundation of Kinetic theory, Proc. Third Berkeley Symp. on Math. Stat. and Prob. 3, 171-197.

[5] H. P. McKean, Jr., A class of Markov processes associated with nonlinear parabolic equations, Proc. Nat. Acad. Sci. 56(1966), 1907-1911.

[6] H. P. McKean, Jr., An exponential formula for solving Boltzmann's equation for a Maxwellian gas, J. Combinatorial Theory 2(1967), 358-382.

[7] H. P. McKean, Jr., Propagation of chaos for a class of non-linear parabolic equations, Lecture series in Differential Equations, session 7, Catholic Univ. (1967).

[8] A. Ya. Povzner, On Boltzmann's equation in the kinetic theory of gases, Mat. Sb. 58(1962), 63-86.

[9] Y. Takahashi, Markov semigroups with simplest interactions I, II (to appear in Proc. Japan Acad.).

[10] H. Tanaka, Propagation of chaos for certain purely discontinuous Markov processes with interactions, J. Fac. Sci., Univ. of Tokyo, Sec. I 17(1970), 259-272.

[11] T. Ueno, A class of Markov processes with interactions I, Proc. Japan Acad. 45(1969), 641-646 ; II, ibd., 995-1000.

[12] T. Ueno, A path space and the propagation of chaos for a Boltzmann's gas model, Proc. Japan Acad., 47(1971), 529-533.

[13] G. E. Uhlenbeck and G. W. Ford, Lectures in Statistical Mechanics, Amer. Math. Soc. Providence 1963.

[14] E. Wild, On Boltzmann's equation in the kinetic theory of gases, Proc. Camb. Phil. Soc. 47(1951), 602-609.

Department of Mathematics
Hiroshima University.

Reprinted from
Proc. 2nd Japan-USSR Symp. on Probab. Th., 478-489,
Lecture Notes in Math., **330**, Springer-Verlag, 1973.

Proc. of Intern. Symp. SDE
Kyoto 1976, pp. 409-425

On the Uniqueness of Markov Process Associated with the Boltzmann Equation of Maxwellian Molecules

Hiroshi TANAKA

§ 1. Introduction

The spatially homogeneous Boltzmann equation of Maxwellian molecules takes the following form:

$$\frac{\partial u(t,x)}{\partial t} = \int (u'u'_1 - uu_1)Q(d\theta)d\varepsilon dx_1 , \tag{1.1}$$

where $Q(d\theta) = Q_M(\theta)\sin\theta d\theta$ with a positive decreasing function $Q_M(\theta)$ such that $Q_M(\theta) \sim \text{const.} \times \theta^{-5/2}$, $\theta \downarrow 0$, and the integration is carried out over $(0,\pi) \times (0,2\pi) \times R^3$. In this paper we consider the following weak version of (1.1):

$$\frac{d}{dt}\langle u,\varphi\rangle = \langle u \otimes u, K\varphi\rangle , \qquad \varphi \in C_0^\infty(R^3) , \tag{1.2}$$

where $C_0^\infty(R^3)$ is the space of real valued C^∞-functions on R^3 with compact supports and $(K\varphi)(x,x_1) = \int \{\varphi(x') - \varphi(x)\}Q(d\theta)d\varepsilon$; x' is on the sphere S_{x,x_1} with center $(x+x_1)/2$ and diameter $|x-x_1|$; θ and ε are the colatitude and the longitude of x', respectively, with respect to a spherical coordinate system on S_{x,x_1} having $x - x_1$ for the polar axis, and finally $u = u(t,\cdot)$ is a probability measure solution to be found.

McKean [3] introduced a class of Markov processes associated with certain nonlinear parabolic equations. The existence problem for (1.1) or (1.2) is almost equivalent to the problem of constructing an associated Markov process of the type introduced by McKean, but it seems that no rigorous results on the existence problem for (1.1) or (1.2), from (non-probabilistic) analysis, have been obtained. The difficulty lies in that $\int_0^\pi Q(d\theta)$ diverges. In [4] I have constructed a Markov process associated with (1.2) by solving the stochastic differential equation:

$$X_t = X_0 + \int_{(0,t]\times S} a(X_s, Y_s(\alpha), \theta, \varepsilon)\hat{p}(dsd\theta d\varepsilon d\alpha) , \tag{1.3a}$$

where $S = (0, \pi) \times (0, 2\pi) \times (0, 1)$, $a(x, x_1, \theta, \varepsilon) = x' - x$ and

(i) $\hat{p}(dsd\theta d\varepsilon d\alpha)$ is a Poisson random measure on $(0, \infty) \times S$ associated with $dsQ(d\theta)d\varepsilon d\alpha$,

(ii) $\{Y_t(\alpha),\ t \geq 0\}$ is a stochastic process defined on the probability space $\{(0, 1), d\alpha\}$, and is equivalent in law to the process $\{X_t, t \geq 0\}$ to be found.

An alternative form of (1.3a) is

$$X_t = X_0 + \sum_{s \leq t} a(s, X_s, \hat{Z}_s) \ , \tag{1.3b}$$

where $a(s, x, \sigma) = a(x, Y_s(\alpha), \theta, \varepsilon)$ for $\sigma = (\theta, \varepsilon, \alpha)$ and $\{\hat{Z}_t, t > 0\}$ is the Poisson point process on S corresponding to the Poisson random measure $\hat{p}$, i.e., defined by the relation $\hat{p}(A) = \sum \chi_A(s, \hat{Z}_s)$ for $A \in \mathscr{B}(R_+ \times S)$. It has been also proved in [4] that the uniqueness in the law sense for solutions of (1.3) holds, however this uniqueness does not mean the uniqueness of the associated Markov process.

The purpose of this paper is to fill this gap by showing that *path functions of any Markov process associated with* (1.2) *can be represented as solutions of* (1.3) *after a suitable extension of basic probability space.* Our task is to find a Poisson point process $\{\hat{Z}_t,\ t > 0\}$ on S with characteristic measure $Q(d\theta)d\varepsilon d\alpha$ such that (1.3b) holds. For this we have to find the compensator of the point process $\{Z_t,\ t > 0\}$ on $R^3 - \{0\}$ defined by $Z_t = X_t - X_{t-}$; this will be done in § 3 and it will follow that $X_t = X_0 + \sum_{s \leq t} Z_s$. Once the compensator of $\{Z_t\}$ is found, a Poisson point process $\{\hat{Z}_t\}$ will be obtained by an application of general results on point processes due to Grigelionis [1] and Karoui-Lepeltier [2]; we will give a construction of $\{\hat{Z}_t\}$ in a simplified form adapted to the present situation.

§ 2. Markov process associated with (1.2)

We begin with the definition of transition function associated with (1.2). Denote by $\mathscr{P}$ the family of probability distributions f on R^3 satisfying $\int_{R^3} |x|\, f(dx) < \infty$, and for $f \in \mathscr{P}$ we put

$$(K_f\varphi)(x) = \int_{R^3} (K\varphi)(x, x_1) f(dx_1) \ .$$

Definition. $\{e_f(t, x, \cdot): f \in \mathscr{P},\ t \geq 0,\ x \in R^3\}$ is called a *transition function associated with* (1.2), if the following five conditions are satisfied.

(e.1) For fixed $f \in \mathscr{P}$, $t \geq 0$ and $x \in R^3$, $e_f(t, x, \cdot)$ is a probability measure on R^3.

(e.2) For fixed $A \in \mathscr{B}(R^3)$, $e_f(t, x, A)$ is jointly measurable in $(f, t, x) \in \mathscr{P} \times R_+ \times R^3$.

(e.3) For each $t \geq 0$ and $f \in \mathscr{P}$, there exists a constant c depending only upon t and f such that

$$\int_{R^3} |y|\, e_f(s, x, dy) \leq c(1 + |x|) , \qquad 0 \leq s \leq t , \quad x \in R^3 .$$

(e.4) If we set $u(t, \cdot) = \int_{R^3} f(dx) e_f(t, x, \cdot)$, then

$$\langle e_f(t, x, \cdot), \varphi \rangle = \varphi(x) + \int_0^t \langle e_f(s, x, \cdot),\ K_{u(s)}\varphi \rangle ds , \qquad \varphi \in C_0^\infty(R^3) .$$

(e.5) (Kolmogorov-Chapman equation)

$$e_f(t, x, \cdot) = \int_{R^3} e_f(s, x, dy) e_{u(s)}(t - s, y, \cdot) , \qquad 0 \leq s \leq t ,$$

where u is the same as in (e.4).

Given a transition function $\{e_f(t, x, \cdot)\}$ associated with (1.2), we can construct a Markov process on R^3. To be precise, let Ω be the space of R^3-valued functions on R_+, and denote the value $\omega(t)$ of $\omega(\in \Omega)$ at t by $X_t(\omega)$ or X_t. We put $\mathscr{B} = \sigma\{X_t : t < \infty\}$ and $\mathscr{B}_t = \sigma\{X_s : s \leq t\}$, where $\sigma\{—\}$ denotes the smallest σ-field on Ω that makes $\{—\}$ measurable. Then there exists a unique family $\{P_f^x : x \in R^3, f \in \mathscr{P}\}$ of probability measures on $(\Omega, \mathscr{B})$ with the following properties (i) and (ii).

(i) For fixed $A \in \mathscr{B}$, $P_f^x(A)$ is jointly measurable in $(x, f) \in R^3 \times \mathscr{P}$.

(ii) For $0 < t_1 < \cdots < t_n$ and $A_k \in \mathscr{B}(R^3)$, $0 \leq k \leq n$,

$$\begin{aligned} &P_f^x\{X_0 \in A_0,\ X_{t_1} \in A_1, \cdots, X_{t_n} \in A_n\} \\ &\quad = \delta_x(A_0) \int_{A_1} e_f(t_1, x, dx_1) \int_{A_2} e_{u(t_1)}(t_2 - t_1, x_1, dx_2) \\ &\qquad \cdots \int_{A_n} e_{u(t_{n-1})}(t_n - t_{n-1}, x_{n-1}, dx_n) . \end{aligned}$$

We put $P_f(\cdot) = \int_{R^3} f(dx) P_f^x(\cdot)$, and denote by E_f (or E_f^x) the expectation with respect to P_f (or P_f^x). Then the following Markov property is proved by a routine method. For a nonegative $\mathscr{B}$-measurable function Φ on Ω we have

$$(2.1a)\qquad E_f^x\{\Phi(\Theta_t\omega)\,|\,\mathscr{B}_t\} = E_{u(t)}^{X_t}\{\Phi\}\,, \qquad P_f^x\text{-a.s.}\,,$$

$$(2.1b)\qquad E_f\{\Phi(\Theta_t\omega)\,|\,\mathscr{B}_t\} = E_{u(t)}^{X_t}\{\Phi\}\,, \qquad P_f\text{-a.s.}\,,$$

where $\Theta_t\colon \Omega \to \Omega$ is the shift operator defined by $X_s(\Theta_t\omega) = X_{t+s}(\omega)$ for $0 \le s < \infty$.

Lemma 1. *If $\varphi\colon R^3 \to R$ is Lipschitz continuous, then*

(i) *there exists a constant c_1 depending only upon the Lipschitz constant of φ such that $|(K\varphi)(x, x_1)| \le c_1|x - x_1|$,*

(ii) *for each $t \ge 0$ and $f \in \mathscr{P}$, there exists a constant c_2 depending only upon t, f and the Lipschitz constant of φ such that*

$$|(K_{u(s)}\varphi)(x)| \le c_2(1 + |x|)\,, \qquad 0 \le \forall s \le t\,,$$

where $u(s) = \int_{R^3} f(dx)e_f(s, x, \cdot)$.

Proof. From $|x' - x| \le \theta|x - x_1|/2$, we have $|\varphi(x') - \varphi(x)| \le \text{const.}\,|x - x_1|\theta$ and hence

$$|(K\varphi)(x, x_1)| \le \text{const.}\,|x - x_1| \int_{(0,\pi)\times(0,2\pi)} \theta Q(d\theta)d\varepsilon = c_1|x - x_1|\,,$$

proving (i). (ii) follows from (i) and (e.3).

The following lemma is an immediate consequence of the above lemma and (e.4),

Lemma 2. *If $\varphi\colon R^3 \to R$ is Lipschitz continuous, then*

$$\langle e_f(t, x, \cdot), \varphi\rangle = \varphi(x) + \int_0^t \langle e_f(s, x, \cdot), K_{u(s)}\varphi\rangle ds\,.$$

Lemma 3. $E_f^x\{|X_t - X_s|\} \le \text{const.}\,(1 + |x|)\,|t - s|$ *for* $0 \le s$, $t \le T$, *where* const. *may depend upon f and T but not upon x.*

Proof. For $0 < s < t$ we have

$$E_f^x\{|x_t - x_s|\} = \int_{R^3} e_f(s, x, dy) \int_{R^3} e_{u(s)}(t - s, y, dz)\,|z - y|\,.$$

We now put $\varphi(z) = |z - y|$ for fixed y and apply Lemma 2:

$$\int_{R^3} e_{u(s)}(t - s, y, dz)\,|z - y|$$
$$= \varphi(y) + \int_0^{t-s} \langle e_{u(s)}(\tau, y, \cdot), K_{u(s-\tau)}\varphi\rangle d\tau$$

$$\leq \int_0^{t-s} \langle e_{u(s)}(\tau, y, \cdot), |K_{u(s+\tau)}\varphi|\rangle d\tau$$
$$\leq c_2 \int_0^{t-s} d\tau \int_{R^3} e_{u(s)}(\tau, y, dz)(1 + |z|) \ .$$

Therefore we have

$$E_f^x\{|X_t - X_s|\} \leq c_2 \int_0^{t-s} d\tau \int_{R^3} e_f(s, x, dy) \int_{R^3} e_{u(s)}(\tau, y, dz)(1 + |z|)$$
$$= c_2 \int_0^{t-s} d\tau \int_{R^3} e_f(s + \tau, x, dz)(1 + |z|)$$
$$\leq \text{const.}\ (1 + |x|)(t - s) \ .$$

Theorem 1. *The stochastic process $\{X_t, P_f^x\}$ has a modification which is right continuous and of bounded variation on each finite t-interval.*

Proof. Let Y_t be any component of X_t. Let $Q_+ = \{r_k\}_{k\geq 1}$ be the set of nonnegative rational numbers, and for any partition of $[0, t]$:

$$\varDelta: 0 = t_0 < t_1 < \cdots < t_n = t \ , \qquad t_j \in Q_+ \quad (0 \leq j \leq n) \ ,$$

we put

$$V_t^\varDelta = \sum_{j=1}^n |Y_{t_j} - Y_{t_{j-1}}| \ ,$$
$$U_t^\varDelta = V_t^\varDelta - (Y_t - Y_0) = 2 \sum_{k=1}^n (Y_{t_j} - Y_{t_{j-1}})^- \ ,$$

and then

$$V_t = \sup V_t^\varDelta \ , \quad U_t = \sup_\varDelta U_t^\varDelta \ , \quad t \in Q_+ \ ,$$

where the supremum is taken over all such partitions $\varDelta$. For each $t \in Q_+$ and $k \geq 1$ we write $\{r_1, \cdots, r_k\} \cap [0, t] = \{\tau_j\}_{1\leq j\leq n}$ with $0 = \tau_0 < \cdots < \tau_n = t$ and then denote by $\varDelta_k$ the partition of $[0, t]$ with partitioning points $\{\tau_j\}_{1\leq j\leq n}$. Then V_t is clearly the increasing limit of $V_t^{\varDelta_k}$ as $k \uparrow \infty$, and hence

$$E_f^x\{V_t\} = \lim_{k\to\infty} E_f^x\{V_t^{\varDelta_k}\} = \lim_{k\to\infty} \sum_{j=1}^n E_f^x\{|Y_{\tau_j} - Y_{\tau_{j-1}}|\}$$
$$\leq \text{const.}\ (1 + |x|)t < \infty \ , \qquad t \in Q_+ \ ,$$

and similarly $E_f^x\{U_t\} < \infty$, $t \in Q_+$. On the other hand, since V_t and U_t

are non-decreasing in $t \in Q_+$, $Y_t = Y_0 + V_t - U_t$, $t \in Q_+$, has right hand limits. Therefore the limit

$$\tilde{X}_t = \lim_{\substack{s \downarrow t \\ s \in Q_+}} X_s , \qquad t \in R_+$$

exists and gives a desired modification of X_t.

By virtue of Theorem 1, the probability measures P_f^π and P_f can be constructed on *the space of R^3-valued right continuous functions on R_+ having bounded variation on each finite t-interval*. From now on, Ω denotes this restricted space and so the σ-fields $\mathscr{B}$ and $\mathscr{B}_t$ are the ones on this new Ω. We call $\{\Omega, \mathscr{B}, X_t, P_f : f \in \mathscr{P}\}$ *a Markov process associated with* (1.2).

§ 3. Point process $\{Z_t\}$

3.1. In general suppose we are given a probability space $(\Omega, \mathscr{F}, P)$, a Borel subset S of R^d and an extra point ∂ not belonging to S. An $S \cup \{\partial\}$-valued process $\{Z_t(\omega), t > 0\}$ defined on $(\Omega, \mathscr{F}, P)$ is called *a point process on S*, if i) $Z_t(\omega)$ is jointly measurable in (t, ω), and ii) $\{t : Z_t(\omega) \in S\}$ is a countable set.

Given a point process $\{Z_t, t > 0\}$ on S, we put

$$p(A) = \sum_t \chi_A(t, Z_t) , \qquad A \in \mathscr{B}((0, \infty) \times S) ,$$

$$p_t(B) = \sum_{0 < s \leq t} \chi_B(Z_s) , \qquad B \in \mathscr{B}(S) , \quad t \geq 0 .$$

Given also a σ-finite Borel measure λ on S, we introduce the following definition. A point process $\{Z_t, t > 0\}$ is called a *Poisson point process on S with characteristic measure λ*, if for any *disjoint* family $A_1, \cdots, A_n$ of Borel sets in $(0, \infty) \times S$ such that $\tilde{\lambda}(A_k) = \int_{A_k} dt d\lambda < \infty$ $(1 \leq k \leq n)$ we have

$$P\{p(A_k) = m_k, k = 1, \cdots, n\} = \prod_{k=1}^n \left\{ e^{-\tilde{\lambda}(A_k)} \frac{(\tilde{\lambda}(A_k))^{m_k}}{m_k !} \right\}$$

for $m_1, \cdots, m_n \in N$.

Then the following characterization of Poisson point processes is well-known.

Theorem 2. *Suppose we are given a point process $\{Z_t, t > 0\}$ on S, a σ-finite Borel measure λ on S and an increasing family $\{\mathscr{F}_t\}$ of sub-σ-fields of $\mathscr{F}$. If, for each $B \in \mathscr{B}(S)$ with $\lambda(B) < \infty$, $\{p_t(B) - \lambda(B)t$,*

$t \geq 0\}$ is an $\{\mathscr{F}_t\}$-martingale, then $\{Z_t,\ t > 0\}$ is a Poisson point process on S with characteristic measure λ.

In the above case, we also call $\{Z_t,\ t > 0\}$ an $\{\mathscr{F}_t\}$-adapted Poisson point process.

3.2. Suppose we are given a Markov process $\{\Omega, \mathscr{B}, X_t, P_f : f \in \mathscr{P}\}$ introduced in § 2, and put $Z_t = X_t - X_{t-}$. Then $\{Z_t,\ t > 0\}$ is a point process on $R_0^3 = R^3 - \{0\}$. The purpose of this subsection is to find the compensator of $\{Z_t\}$.

Fixing $f \in \mathscr{P}$, we introduce the following notations.

$$\mathscr{F} = \text{the completion of } \mathscr{B} \text{ with respect to } P_f\ ,$$
$$\mathscr{F}_t^0 = \{A \in \mathscr{F} : P_f(A \ominus B) = 0 \text{ for some } B \in \mathscr{B}_t\}\ ,$$
$$\mathscr{F}_t = \textstyle\bigcap_{\varepsilon>0} \mathscr{F}_{t+\varepsilon}^0\ .$$

We now think of the unit interval $(0, 1)$ as a probability space by considering the Lebesgue measure on $\mathscr{B}((0, 1))$, and we take an R^3-valued stochastic process $\{Y_t(\alpha),\ t \geq 0\}$ defined on this probability space $(0, 1)$ with the following properties: (a) Sample paths of $\{Y_t\}$ are right continuous and have bounded variation on each finite t-interval, and (b) $\{Y_t,\ t \geq 0\}$ is equivalent in law to the process $\{X_t,\ t \geq 0,\ P_f\}$. Next, we put $S = (0, \pi) \times (0, 2\pi) \times (0, 1)$ and denote by $\sigma = (\theta, \varepsilon, \alpha)$ a generic element of S. On S we consider the Borel measure λ defined by $d\lambda = Q(d\theta) \otimes d\varepsilon \otimes d\alpha$. We also put

$$a(x, x_1, \theta, \theta, \varepsilon) = x' - x\ ,$$
$$a(t, x, \sigma) = a(x, Y_t(\alpha), \theta, \varepsilon) \qquad \text{for } \sigma = (\theta, \varepsilon, \alpha)\ ,$$
$$n(x, x_1, A) = \int_{(0,\pi)\times(0,2\pi)} \delta_{x'-x}(A) Q(d\theta) d\varepsilon\ , \qquad A \in \mathscr{B}(R_0^3)\ ,$$
$$n_u(x, A) = \int_{R^3} n(x, x_1, A) u(dx_1)\ , \qquad A \in \mathscr{B}(R_0^3)\ .$$

Then, $\langle n_{u(t)}(y, \cdot), \psi \rangle = \langle \lambda, \psi(a(t, y, \cdot)) \rangle$ holds for any non-negative Borel function ψ on R^3 with $\psi(0) = 0$, where $u(t, \cdot) = \int_{R^3} f(dx) e_f(t, x, \cdot)$. Finally we put

$$p_t(\varphi) = \sum_{s \leq t} \varphi(s, X_{s-}, Z_s)$$

for a real valued Borel function φ on $R_+ \times R^3 \times R^3$ with $\varphi(t, y, 0) = 0$.

Theorem 3. *Let $\varphi : R_+ \times R^3 \times R^3 \to R$ be a Borel function with $\varphi(t, y, 0) = 0$ and assume that*

$$(3.1)\qquad \int_0^t ds \int_{R^3} u(s, dy) \int_{R_0^3} n_{u(s)}(y, dz)\,|\varphi(s, y, z)| < \infty$$

for each $t \in R_+$. *Then*

$$p_t(\varphi) - \int_0^t \langle \lambda, \varphi(s, X_s, a(s, X_s, \cdot))\rangle ds$$

is an $\{\mathscr{F}_t\}$*-martingale with respect to* P_f. *In other words*

$$\int_0^t \langle \lambda, \varphi(s, X_s, a(s, X_s, \cdot))\rangle ds$$

is the compensator of $p_t(\varphi)$.

Remark. The condition (3.1) is satisfied, if $|\varphi(t, y, z)| \leq$ const. $|z|$.

Before going to the proof of this theorem, we prepare some lemmas. First we assume that $\varphi \in C_0^\infty(R_+ \times R^3 \times R_0^3)$. For fixed $t > 0$ we consider a partition of $[0, t]$:

$$\Delta : 0 = t_0 < t_1 < \cdots < t_n = t\,,$$

and put

$$I_\Delta = E_f^x\left\{\sum_{k=1}^n \varphi(t_{k-1}, X_{t_{k-1}}, X_{t_k} - X_{t_{k-1}})\right\}\,.$$

Then, with the notation $\varphi^{k,y}(z) = \varphi(t_{k-1}, y, z - y)$ we have

$$\begin{aligned} I_\Delta &= \sum_{k=1}^n \int_{R^3} e_f(t_{k-1}, x, dy) \int_{R^3} e_{u(t_{k-1})}(t_k - t_{k-1}, y, dz)\varphi^{k,y}(z) \\ &= \sum_{k=1}^n \int_{R^3} e_f(t_{k-1}, x, dy) \int_0^{t_k - t_{k-1}} ds \int_{R^3} e_{u(t_{k-1})}(s, y, dz) K_{u(t_{k-1}+s)}\varphi^{k,y}(z)\,, \\ &= I'_\Delta + I''_\Delta\,, \end{aligned}$$

where

$$\begin{aligned} I'_\Delta = \sum_{k=1}^n \int_{R^3} e_f(t_{k-1}, x, dy) \int_0^{t_k - t_{k-1}} ds \int_{R^3} e_{u(t_{k-1})}(s, y, dz) \\ \times \chi_\varepsilon(z - y) K_{u(t_{k-1}+s)}\varphi^{k,y}(z)\,, \end{aligned}$$

and I''_Δ denotes a similar formula obtained from I'_Δ with the replacement of

$$\chi_\varepsilon(z - y) = \begin{cases} 1 & \text{for } |z - y| > \varepsilon \\ 0 & \text{for } |z - y| \leq \varepsilon \end{cases}$$

by $\bar{\chi}_\varepsilon(z-y) = 1 - \chi_\varepsilon(z-y)$.

Lemma 4. *If* $\varphi \in C_0^\infty(R_+ \times R^3 \times R_0^3)$, *then for any* $\varepsilon > 0$

$$I'_\Delta \to 0 \quad as \quad |\Delta| = \max_{1 \le k \le n} (t_k - t_{k-1}) \to 0 .$$

Proof. By Lemma 1 there exists a constant c independent of y, z, k and s such that

$$|K_{u(t_{k-1}+s)}\varphi^{k,y}(z)| \le c(1 + |z|) .$$

If we put

$$\rho(z) = \begin{cases} 0 & \text{for } |z| \le \varepsilon/2 \\ 2|z|/\varepsilon - 1 & \text{for } \varepsilon/2 < |z| < \varepsilon \\ 1 & \text{for } |z| \ge \varepsilon , \end{cases}$$

$$\xi^y(z) = c\rho(y - z)(1 + |z|) ,$$

then we have

$$\chi_\varepsilon(z) \le \rho(z) ,$$

$$|I'_\Delta| \le \sum_{k=1}^n \int_{R^3} e_f(t_{k-1}, x, dy) \int_0^{t_k - t_{k-1}} ds \int_{R^3} e_{u(t_{k-1})}(s, y, dz)\xi^y(z) .$$

Since the support of φ is compact, the integral with respect to $e_{u(t_{k-1})}(s, y, dz)$ in the above may be performed only on some compact set B of R^3. Moreover, $\xi^y(z)$ is Lipschitz continuous as a function of z for each fixed y, and the Lipschitz constant is bounded as far as y is on B. Therefore, by Lemma 2 we have

$$\begin{aligned} \int_{R^3} e_{u(t_{k-1})}(s, y, dz)\xi^y(z) &= \int_0^s d\tau \int_{R^3} e_{u(t_{k-1})}(\tau, y, dz) K_{u(t_{k-1}+\tau)}\xi^y(z) \\ &\le \text{const.} \int_0^s d\tau \int_{R^3} e_{u(t_{k-1})}(\tau, y, dz)(1 + |z|) , \end{aligned}$$

and hence

$$\begin{aligned} |I'_\Delta| &\le \text{const.} \sum_{k=1}^n \int_{R^3} e_f(t_{k-1}, x, dy) \int_0^{t_k - t_{k-1}} ds \int_0^s d\tau \\ &\qquad \times \int_{R^3} e_{u(t_{k-1})}(\tau, y, dz)(1 + |z|) \\ &= \text{const.} \sum_{k=1}^n \int_0^{t_k - t_{k-1}} ds \int_0^s d\tau \int_{R^3} e_f(t_{k-1} + \tau, x, dz)(1 + |z|) \end{aligned}$$

$$\leq \text{const.} \sum_{k=1}^{n} (t_k - t_{k-1})^2 \to 0 \,, \qquad \text{as } |\Delta| \to 0 \,.$$

Lemma 5. *If $\varphi \in C_0^\infty(R_+ \times R^3 \times R_0^3)$, then*

$$\lim_{|\Delta|\to 0} I_\Delta = E_f^x\left\{\int_0^t \langle n_{u(s)}(X_s, \cdot), \varphi(s, X_s, \cdot)\rangle ds\right\} .$$

Proof. Define $\Delta(s)$, $0 \leq s \leq t$, by

$$\Delta(0) = 0 \,, \quad \Delta(s) = t_{k-1} \qquad \text{for } t_{k-1} < s \leq t_k \ (1 \leq k \leq n) \,,$$

and for $\varepsilon > 0$ (small enough) put

$$\begin{aligned} \Phi(s, y, z, z_1) &= (K\varphi^{k,y})(z, z_1)\bar{\chi}_\varepsilon(z - y) \\ &= \int_{R_0^3} \varphi(s, y, z - y + w)\bar{\chi}_\varepsilon(z - y) n(z, z_1, dw) \,. \end{aligned}$$

Then

$$\begin{aligned} I''_\Delta &= \sum_{k=1}^{n} \int_{R^3} e_f(t_{k-1}, x, dy) \int_0^{t_k - t_{k-1}} ds \int_{R^3} e_{u(t_{k-1})}(s, y, dz) \\ &\qquad\qquad \times \int_{R^3} u(t_{k-1} + s, dz_1)\Phi(t_{k-1}, y, z, z_1) \\ &= \sum_{k=1}^{n} \int_0^{t_k - t_{k-1}} ds \int_0^1 d\alpha E_f^x\{\Phi(t_{k-1}, X_{t_{k-1}}, X_{t_{k-1}+s}, Y_{t_{k-1}+s}(\alpha))\} \\ &= \int_0^1 d\alpha E_f^x\left\{\int_0^t \Phi(\Delta(s), X_{\Delta(s)}, X_s, Y_s(\alpha)) ds\right\} \\ &= \int_0^1 d\alpha E_f^x\left\{\int_{[0,t]\times R_0^3} ds\, n(x_s, Y_s(\alpha), dz) \right. \\ &\qquad\qquad \left. \times\ \varphi(\Delta(s), X_{\Delta(s)}, X_s - X_{\Delta(s)} + z)\bar{\chi}_\varepsilon(X_s - X_{\Delta(s)})\right\} \\ &\to \int_0^1 d\alpha E_f^x\left\{\int_{[0,t]\times R_0^3} ds\, n(X_{s-}, Y_s(\alpha), dz)\varphi(s, X_{s-}, z)\right\} \,, \quad \text{as } |\Delta| \to 0 \\ &= E_f^x\left\{\int_0^t \langle n_{u(s)}(X_s, \cdot), \varphi(s, X_s, \cdot)\rangle ds\right\} . \end{aligned}$$

This combined with Lemma 4 completes the proof.

Lemma 6. *Let $\varphi: R_+ \times R^3 \times R^3 \to R_+$ be a Borel function with $\varphi(t, y, 0) = 0$. Then*

$$E_f^x\{p_t(\varphi)\} = E_f^x\left\{\int_0^t \langle \lambda, \varphi(s, X_s, a(s, X_s, \cdot))\rangle ds\right\} . \tag{3.2}$$

Proof. We first consider the case $\varphi \in C_0^\infty(R_+ \times R^3 \times R_0^3)$. In this case we have

$$p_t^\Delta(\varphi) \equiv \sum_{k=1}^{n} \varphi(t_{k-1}, X_{t_{k-1}}, X_{t_k} - X_{t_{k-1}}) \to p_t(\varphi) ,$$

as $|\Delta| \to 0$. Since $|\varphi(s, y, z)| \leq$ const. $|z|$, we have also

$$|p_t^\Delta(\varphi)| \leq \text{const.} \sum_{k=1}^{n} |X_{t_k} - X_{t_{k-1}}| \leq V ,$$

where V is some P_f^x-integrable random variable. Therefore by Lebesgue's dominated convergence theorem,

$$\begin{aligned} E_f^x\{p_t(\varphi)\} &= \lim_{|\Delta|\to 0} E_f^x\{p_t^\Delta(\varphi)\} \\ &= E_f^x\left\{\int_0^t \langle n_{u(s)}(X_s, \cdot), \varphi(s, X_s, \cdot)\rangle ds\right\} \\ &= E_f^x\left\{\int_0^t \langle \lambda, \varphi(s, X_s, a(s, X_s, \cdot))\rangle ds\right\} , \end{aligned}$$

and this proves (3.2) for smooth φ. Finally, it is easy to remove the smoothness condition of φ in (3.2).

We now come to the proof of Theorem 3. For $0 \leq s \leq t$ we have

$$\begin{aligned} &p_t(\varphi) = p_s(\varphi) + p_{t-s}(\varphi^s) \circ \Theta_s , \\ &\int_0^t \langle \lambda, \varphi(\tau, X_\tau, a(\tau, X_\tau, \cdot))\rangle d\tau \\ &\quad = \int_0^s \langle \lambda, \varphi(\tau, X_\tau, a(\tau, X_\tau, \cdot))\rangle d\tau \\ &\qquad + \int_0^{t-s} \langle n_{u(s+\tau)}(X_\tau(\Theta_s\omega), \cdot), \varphi^s(\tau, X_\tau(\Theta_s\omega), a(\tau, X_\tau(\Theta_s\omega), \cdot))\rangle d\tau , \end{aligned}$$

where $\varphi^s(\tau, y, z) = \varphi(s + \tau, y, z)$. Therefore, if we put

$$q_t(\varphi) = p_t(\varphi) - \int_0^t \langle \lambda, \varphi(\tau, X_\tau, a(\tau, X_\tau, \cdot))\rangle d\tau ,$$

then we have $q_t(\varphi) = q_s(\varphi) + q_{t-s}(\varphi^s) \circ \Theta_s$, and hence by using the Markov property (2.1a)

$$\begin{aligned} E_f\{q_t(\varphi) | \mathscr{B}_s\} &= q_s(\varphi) + E_f\{q_{t-s}(\varphi^s) \circ \Theta_s | \mathscr{B}_s\} \\ &= q_s(\varphi) + E_{u(s)}^{x_s}\{q_{t-s}(\varphi^s)\} \\ &= q_s(\varphi) , \qquad\qquad P_f\text{-a.s. .} \end{aligned}$$

Thus $\{q_t(\varphi)\}$ is a $\{\mathscr{B}_t\}$-martingale, and hence an $\{\mathscr{F}_t\}$-martingale.

Corollary. $X_t = X_0 + \sum_{s\le t} Z_s$, P_f-a.s. .

Proof. For $\varphi(z) = z$ we have $\sum_{s\le t} Z_s = p_t(\varphi)$, and $E_f\{p_t(\varphi)\} = E_f\{X_t - X_0\}$ by making use of Theorem 3 and Lemma 2. Therefore $X_t - X_0 - \sum_{s\le t} Z_s$ is a martingale with respect to P_f; also it is continuous and has bounded variation on each finite t-interval. But, such a martingale must be 0, and so the corollary is proved.

§ 4. Derivation of stochastic differential equation

Notations are the same as in the subsection 3.2. We now know the compensator of the point process $\{Z_t\}$, and so the representation (1.3b) of $\{Z_t\}$ by means of certain Poisson point process $\{\hat{Z}_t\}$ might be a consequence of general works due to Grigelionis [1] and Karoui-Lepeltier [2]. However we give the construction of $\{\hat{Z}_t\}$ in detail, because we wish the whole proof to be self-contained. The construction given here seems to be simpler.

Theorem 4. *For each fixed* $f \in \mathscr{P}$ *we can find a probability space* $\{\hat{\Omega}, \hat{\mathscr{F}}, \hat{P}\}$, *an increasing family* $\{\hat{\mathscr{F}}_t\}$, *a mapping* π *from* $\hat{\Omega}$ *onto* Ω *and an* $\{\hat{\mathscr{F}}_t\}$*-adapted Poisson point process* $\{\hat{Z}_t, t > 0\}$ *on* S *with characteristic measure* λ, *having the following properties.*

(i) $\pi^{-1}(\mathscr{F}_t) \subset \hat{\mathscr{F}}_t$, $\pi^{-1}(\mathscr{F}) \subset \hat{\mathscr{F}}$ *and* $\hat{P}\{\pi^{-1}(A)\} = P_f\{A\}$ *for* $A \in \mathscr{F}$.

(ii) $X_t \circ \pi = X_0 \circ \pi + \sum_{s\le t} a(s, X_{s-}\circ\pi, \hat{Z}_s)$, $\hat{P}$*-a.s.*,

or what is the same

$$X_t \circ \pi = X_0 \circ \pi + \int_{(0,t]\times S} a(X_{s\cdot}\circ\pi, Y_s(\alpha), \theta, \varepsilon)\hat{p}(dsd\sigma)\ , \qquad \hat{P}\text{-a.s. },$$

where $\hat{p}(A) = \sum \chi_A(t, \hat{Z}_t)$ *for* $A \in \mathscr{B}((0, \infty) \times S)$.

Lemma 7. *There exists* $Q(t, y, z, A)$ *such that*

(i) *for fixed* $t \ge 0$, $y \in R^3$ *and* $z \in R^3_0$, $Q(t, y, z, \cdot)$ *is a probability measure on* S,

(ii) *for fixed* $A \in \mathscr{B}(S)$, $Q(t, y, z, A)$ *is jointly measurable in* (t, y, z),

(iii) *for any nonnegative Borel function* φ *on* $R^3 \times S$ *with* $\varphi(0, \sigma) = 0$ *and for any* $t \ge 0$, $y \in R^3$,

$$\int_S \lambda(d\sigma)\varphi(a(t, y, \sigma), \sigma) = \int_{R^3_0\times S} n_{u(t)}(y, dz)Q(t, y, z, d\sigma)\varphi(z, \sigma)\ .$$

Proof of the lemma. For fixed $t \ge 0$, $y \in R^3$ and $A \in \mathscr{B}(S)$ we put

$$n^A_{u(t)}(y, B) = \int_A \chi_B(a(t, y, \sigma))\lambda(d\sigma)\ , \qquad B \in \mathscr{B}(R^3_0)\ .$$

Then $Q(t, y, z, A)$, as a function of z, should be the Radon-Nikodym derivative of $n^A_{u(t)}(y, \cdot)$ with respect to $n_{u(t)}(y, \cdot)$. A nice version of $Q(t, y, z, A)$ as stated in the lemma can be obtained by making use of the convergence theorem of martingales.

In order to prove Theorem 4 we must prepare some probability spaces together with various quantities defined on them.

1°. $\{\Omega, \mathscr{F}, P\}$: This is, of course, the basic probability space on which our Markov process $\{X_t\}$ has been given. Let

$$B_0 = \{z \in R^3 : |z| > 1\}\ , \qquad B_n = \left\{z \in R^3 : \frac{1}{n+1} < |z| \leq \frac{1}{n}\right\} \qquad n \geq 1\ ,$$

and define $\{\mathscr{F}_t\}$-stopping times T_{nk}, $n \geq 0$, $k \geq 1$, by

$$T_{n0}(\omega) = \inf\{t > 0 : Z_t(\omega) \in B_n\}\ , \qquad n \geq 0$$
$$T_{nk}(\omega) = \inf\{t > T_{n,k-1}(\omega) : Z_t(\omega) \in B_n\}\ , \qquad n \geq 0\ ,\ k \geq 1\ .$$

2°. $\{\Omega', \mathscr{F}', P'\}$: This is the probability space obtained by taking $\Omega' = (0, 1)$, $\mathscr{F}' = \mathscr{B}((0, 1))$ and P' = the Lebesgue measure (restricted on $\mathscr{F}'$). On this probability space we choose a sequence of independent random variables $\{\xi_{nk}(\omega') : n \geq 0,\ k \geq 1\}$, each being uniformly distributed on (0, 1). Moreover we choose a jointly measurable function $Y(t, y, z, \omega')$ on $R_+ \times R^3 \times R^3_0 \times \Omega'$ such that, for each $t \geq 0$, $y \in R^3$ and $z \in R^3_0$, $Y(t, y, z, \cdot)$ is an S-valued random variable defined on $\{\Omega', \mathscr{F}', P'\}$ and with probability distribution $Q(t, y, z, \cdot)$.

3°. $\{\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P}\}$: We choose an arbitrary probability space $\{\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P}\}$ on which there is defined a Poisson point process $\{\tilde{Z}_t, t > 0\}$ on S with characteristic measure λ, and put $\tilde{\mathscr{F}}_t = \sigma\{\tilde{Z}_s : s \leq t\}$.

Now we can construct all that we need.

(a) $\hat{\Omega} = \Omega \times \Omega' \times \tilde{\Omega}\ ,\quad \hat{\mathscr{F}} = \mathscr{F} \otimes \mathscr{F}' \otimes \tilde{\mathscr{F}}\ ,\quad \hat{P} = P_f \otimes P' \otimes \tilde{P}\ .$

(b) $\hat{\mathscr{F}}_t$ = the σ-field on $\hat{\Omega}$ generated by all sets of the form

$$(A \cap B) \times A' \times \tilde{A}\ , \tag{4.1}$$

where $B \in \mathscr{F}_t$, $\tilde{A} \in \tilde{\mathscr{F}}_t$ and

$$A = \{T_{nk} \leq t \text{ for } \forall(n, k) \in M\}\ ,$$
$$A' = \{\xi_{nk} \in B_{nk} \text{ for } \forall(n, k) \in M\}\ ,$$

for some $M \subset N \times N$ and $B_{nk} \in \mathscr{B}((0, 1))$.

(c) $$Z'_t = \begin{cases} Y(T_{nk}, X_{nk}, Z_{nk}, \xi_{nk}) & \text{if } t = T_{nk} \\ 0 & \text{if } t \neq T_{nk} \quad \text{for } \forall(n, k)\ , \end{cases}$$

where $X_{nk} = X_{T_{nk}-}$ and $Z_{nk} = X_{T_{nk}} - X_{T_{nk}-}$.

(d) $$Z''_t = \begin{cases} \tilde{Z}_t & \text{if } a(t, X_{t-}, \tilde{Z}_t) = 0 \\ 0 & \text{otherwise}\ . \end{cases}$$

(e) $$\hat{Z}_t = Z'_t + Z''_t\ .$$

Then, (i) of Theorem 4 is obvious with the self-evident notation π (projection), and the rest will be proved in the following two lemmas.

Lemma 8. *$\{\hat{Z}_t, t > 0\}$ is an $\{\hat{\mathscr{F}}_t\}$-adapted Poisson point process on S with characteristic measure λ.*

Proof. Since the $\{\hat{\mathscr{F}}_t\}$-adaptedness of $\{\hat{Z}_t,\ t > 0\}$ is clear from the definition of $\hat{Z}_t$, it is enough to prove that

$$\hat{E}\{\hat{p}_t(\varphi) - \hat{p}_s(\varphi) \mid \hat{\mathscr{F}}_s\} = (t - s)\langle\lambda, \varphi\rangle\ , \qquad \hat{P}\text{-a.s.}$$

for any λ-integrable Borel function φ on S and $s \leq t$, where $\hat{p}_t(\varphi) = \sum_{s \leq t} \varphi(\hat{Z}_s)$. For this, it is also enough to prove that

$$\int_{\hat{A}} \{\hat{p}_t(\varphi) - \hat{p}_s(\varphi)\} d\hat{P} = (t - s)\langle\lambda, \varphi\rangle \hat{P}\{\hat{A}\}$$

for any set $\hat{A}$ of the form (4.1). Since

$$p'_t(\varphi) - p'_s(\varphi) = \sum_{n,k} \chi_{(s,t]}(T_{nk})\varphi(Y(T_{nk}, X_{nk}, Z_{nk}, \xi_{nk}))\ ,$$

we have

$$\begin{aligned}
(4.2) \quad & \int_{(A \cap B) \times A'} \{p'_t(\varphi) - p_s(\varphi)\} d(P_f \otimes P') \\
&= \int_{(A \cap B) \times A'} \sum_{(n,k) \notin M} \chi_{(s,t]}(T_{nk})\varphi(Y(T_{nk}, X_{nk}, Z_{nk}, \xi_{nk})) d(P_f \otimes P') \\
&= P'\{A'\} \sum_{(n,k) \notin M} \int_{(A \cap B) \times \Omega'} \chi_{(s,t]}(T_{nk})\varphi(Y_{nk}, X_{nk}, Z_{nk}, \xi_{nk})) d(P_f \otimes P') \\
&= P'\{A'\} \int_0^1 d\alpha \int_{A \cap B} \sum_{s < \tau \leq t} \chi(Z_\tau)\varphi(Y(\tau, X_{\tau-}, Z_\tau, \alpha)) dP_f{}^{*)} \\
&= P'\{A'\} \int_0^1 d\alpha \int_{(A \cap B) \times (s,t] \times R_0^8} dP_f d\tau n_{u(\tau)}(X_{\tau-}, dz)\varphi(Y(\tau, X_{\tau-}, z, \alpha))
\end{aligned}$$

*) $\chi(z) \neq 1$ for $z \neq 0$, and $=0$ for $z = 0$.

$$= \int_{(A\cap B)\times A'} d(P_f \otimes P') \int_{(s,t]\times R_0^3\times S} d\tau n_{u(\tau)}(X_{\tau-}, dz) Q(\tau, X_{\tau-}, z, d\sigma)\chi(z)\varphi(\sigma)$$

$$= \int_{(A\cap B)\times A'} d(P_f \otimes P') \int_{(s,t]\times S} d\tau\lambda(d\sigma)\chi(a(\tau, X_{\tau-}, \sigma))\varphi(\sigma)\,;$$

in the above we used Theorem 3 and Lemma 7. We have also

$$\begin{aligned}(4.3) \qquad & \int_{\hat{A}} \{p_t''(\varphi) - p_s''(\varphi)\} d\hat{P} \\ &= \int_{(A\cap B)\times A'} d(P_f \otimes P') \int_{\tilde{A}} \{p_t''(\varphi) - p_s''(\varphi)\} d\tilde{P} \\ &= \int_{(A\cap B)\times A'} d(P_f \otimes P') \int_{\tilde{A}} d\tilde{P} \int_{(s,t]\times S} d\tau\lambda(d\sigma) \\ &\qquad \times \{1 - \chi(a(\tau, X_{\tau-}, \sigma))\}\varphi(\sigma) \\ &= \int_{\hat{A}} d\hat{P} \int_{(s,t]\times S} d\tau\lambda(d\sigma)\{1 - \chi(a(\tau, X_{\tau-}, \sigma))\}\varphi(\sigma)\,.\end{aligned}$$

Combining (4.2) with (4.3), we obtain

$$\begin{aligned}\int_{\hat{A}} \{\hat{p}_t(\varphi) - \hat{p}_s(\varphi)\} d\hat{P} &= \tilde{P}\{\tilde{A}\} \int_{(A\cap B)\times A'} \{p_t'(\varphi) - p_s'(\varphi)\} d(P_f \otimes P') \\ &\quad + \int_{\hat{A}} \{p_t''(\varphi) - p_s''(\varphi)\} d\hat{P} \\ &= (t-s)\langle\lambda, \varphi\rangle \hat{P}\{\hat{A}\}\,,\end{aligned}$$

as was to be proved.

Lemma 9. *$Z_t \circ \pi = a(t, X_{t-}, \hat{Z}_t)$, $\hat{P}$-a.s., in which we have put $a(t, y, 0) = 0$.*

Proof. Since the sets $\{t\colon Z_t' \in S\}$ and $\{t\colon Z_t'' \in S\}$ are disjoint with $\hat{P}$-probability 1, we have $a(t, X_{t-}, \hat{Z}_t) = a(t, X_{t-}, Z_t') + a(t, X_{t-}, Z_t'') = a(t, X_{t-}, Z_t')$ with $\hat{P}$-probability 1, and hence it is enough to prove that for each fixed n and k

$$Z_{nk} = a(T_{nk}, X_{nk}, Y(T_{nk}, X_{nk}, Z_{nk}, \xi_{nk}))\,, \qquad \hat{P}\text{-a.s.}\,.$$

Putting

$$\Phi^{\omega'}(t, y, z) = \chi_{B_n}(z)\,|z - a(t, y, Y(t, y, z, \xi_{nk}(\omega')))|\,,$$

and then applying Theorem 3 and Lemma 7, we have

$$\hat{E}\{\chi_{(0,t]}(T_{nk})\,|Z_{nk}-a(T_{nk},X_{nk},Y(T_{nk},X_{nk},Z_{nk},\xi_{nk}))|\}$$
$$\le \hat{E}\Big\{\sum_{s\le T_{nk}\wedge t}\Phi^{\omega'}(s,X_{s-},Z_s)\Big\}$$
$$\le E'\Big\{\int_{[0,t]\times R^3\times R^3_0} dsu(s,dy)n_{u(s)}(y,dz)\Phi^{\omega'}(s,y,z)\Big\}$$
$$=\int_{[0,t]\times R^3\times R^3_0\times S} dsu(s,dy)n_{u(s)}(y,dz)Q(s,y,z,d\sigma)\chi_{B_n}(z)\,|z-a(s,y,\sigma)|$$
$$=\int_{[0,t]\times R^3\times S} dsu(s,dy)\lambda(d\sigma)\chi_{B_n}(a(s,y,\sigma))\,|a(s,y,\sigma)-a(s,y,\sigma)|=0\ .$$

as required.

Combining the above two lemmas with the corollary to Theorem 3, we now complete the proof of Theorem 4.

§ 5. Concluding remark

The existence and uniqueness of solutions for the stochastic differential equation (1.3) were discussed by Tanaka [4]. The main results of [4] read as follows.

(A) Let $f\in\mathscr{P}$ be fixed. Then, on a suitable probability space $\{\hat{\Omega},\hat{\mathscr{F}},\hat{P}\}$ with an increasing family $\{\hat{\mathscr{F}}_t\}$ of sub-σ-fields we can construct an $\{\hat{\mathscr{F}}_t\}$-adapted Poisson point process $\{\hat{Z}_t,\ t>0\}$ on S with characteristic measure λ and an $\{\hat{\mathscr{F}}_t\}$-adapted right continuous process $\{X_t,\ t\ge 0\}$ on R^3 with initial distribution f such that

(i) $\int_0^t \hat{E}\{|X_s|\}ds<\infty$ for each $t<\infty$,

(ii) (1.3b) holds with probability 1,

where $\{Y_t(\alpha),\ t\ge 0\}$ is some right continuous process defined on the probability space $\{(0,1),\ d\alpha\}$ and is equivalent in law to the solution process $\{X_t,\ t\ge 0\}$. Moreover, the probability measure on the path space induced by $\{X_t,\ t\ge 0\}$ is uniquely determined from f.

(B) Let $f\in\mathscr{P}$ and $x\in R^3$ be fixed. Then, we can construct a Poisson point process $\{\hat{Z}_t,\ t>0\}$ on S with characteristic measure λ and an $\{\hat{\mathscr{F}}_t\}$-adapted right continuous process $\{X_t^x,\ t\ge 0\}$ on R^3 such that

(iii) $\int_0^t \hat{E}\{|X_s^x|\}ds<\infty$ for each $t<\infty$,

(iv) $X_t^x=x+\sum_{s\le t}a(s,X_{s-}^x,\hat{Z}_s)$, a.s.,

where $a(s,y,\sigma)$ is the same as used in (1.3b). Moreover, the probability measure on the path space induced by $\{X_t^x,\ t\ge 0\}$ is uniquely determined by f and x, and

$$e_f(t,x,A)=\hat{P}\{X_t^x\in A\}\ ,\qquad A\in\mathscr{B}(R^3)\ ,$$

gives a transition function associated with (1.2) in the sense of § 2.

Combining these results with Theorem 4, we obtain the existence and *uniqueness* of Markov process associated with (1.2).

References

[1] Grigelionis, B., On representation of integer-valued random measures by means of stochastic integrals with respect to the Poisson measure, Liet. Mat. Sb., XI (1971), 93–108, (Russian).

[2] Karoui, N. E. and Lepeltier, J. P., Représentation des processus ponctuels multivariés à l'aide d'un processus de Poisson, Z. Wahrscheinlichkeits-theorie verw. Gebiete **39** (1977), 111–133.

[3] McKean, H. P., A class of Markov processes associated with non-linear parabolic equations, Proc. Nat. Acad. Sci. U.S.A., **56** (1966), 1907–1911.

[4] Tanaka, H., On Markov process corresponding to Boltzmann's equation of Maxwellian gas, Proc. 2nd Japan-USSR Symp. Prob. Th., Lecture Notes in Mathematics, **330**, Springer, 478–489.

[5] Uhlenbeck, G. E. and Ford, G. W., Lectures in statistical mechanics, Lectures in Appl. Math., vol. **1**, Amer. Math. Soc., Providence, 1963.

Department of Mathematics
Hiroshima University
Hiroshima 730, Japan

Z. Wahrscheinlichkeitstheorie verw. Gebiete
46, 67–105 (1978)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Probabilistic Treatment of the Boltzmann Equation of Maxwellian Molecules

Hiroshi Tanaka

Department of Mathematics, Faculty of Science, Hiroshima University,
J-730 Hiroshima, Japan

Introduction

The Boltzmann equation in the kinetic theory of dilute gases is the equation that governs the time evolution of the number density $u(t,x)$ given by

$$u(t,x)\,dx = \frac{\text{the number of molecules with velocities} \in dx \text{ at time } t}{\text{the total number of molecules}}$$

for a gas composed of a very large number of molecules moving in space according to the law of classical mechanics and colliding in pairs. Here we assume spatial homogeneity. When there is no outside force, the equation is

$$\frac{\partial u}{\partial t} = \int_{(0,\infty)\times(0,2\pi)\times\mathbf{R}^3} (u' u_1' - u u_1)\,|x - x_1|\, r\,dr\,d\varphi\,dx_1, \qquad t \geqq 0,\ x \in \mathbf{R}^3$$

where $u = u(t,x)$, $u_1 = u(t,x_1)$, $u' = u(t,x')$ and $u_1' = u(t,x_1')$. If we denote by S_{x,x_1} the sphere with center $(x+x_1)/2$ and diameter $|x-x_1|$, then x' and x_1' (the velocities of molecules after "collision") are always on the sphere S_{x,x_1} or more precisely $S_{x,x_1} = S_{x',x_1'}$ according to the conservation laws of momentum and energy. We consider a spherical coordinate system with polar axis defined by the relative velocity $x - x_1$, and put

$\theta =$ the colatitude of x'

$\varphi =$ the longitude of x'.

The Maxwellian gas is the case in which molecules repel each other with a force inversely proportional to the fifth power of their distance, and in this case the colatitude θ is determined from the impact parameter r by the following relation:

$$\frac{\pi - \theta}{2} = \int_0^{\rho_0} \frac{d\rho}{\sqrt{1 - \rho^2 - \dfrac{4}{|x - x_1|^2}\, U\left(\dfrac{r}{\rho}\right)}}, \tag{0.1}$$

where $U(\rho)=\text{const}\cdot\rho^{-4}$ and ρ_0 is the positive root of

$$1-\rho^2-\frac{4}{|x-x_1|^2}U\left(\frac{r}{\rho}\right)=0.$$

From the relation (0.1) we have $|x-x_1|\,r\,dr=Q_M(\theta)\sin\theta\,d\theta$ with some positive decreasing function $Q_M(\theta)$ of θ such that $Q_M(\theta)\sim\text{const}\cdot\theta^{-5/2}$ as $\theta\downarrow 0$. Thus we have the following Boltzmann equation of Maxwellian molecules

$$\frac{\partial u}{\partial t}=\int_{(0,\pi)\times(0,2\pi)\times\mathbf{R}^3}(u'u'_1-uu_1)\,Q(d\theta)\,d\varphi\,dx_1, \tag{0.2}$$

where $Q(d\theta)=Q_M(\theta)\sin\theta\,d\theta$. For these matters, see Uhlenbeck and Ford [20]. The fact that $Q(d\theta)$ does not involve $|x-x_1|$ is a consequence of the inverse fifth power force, and in this sense the situation is simplified. But difficulties arise from the non-cutoff type $\left(\int_0^\pi Q(d\theta)=\infty\right)$ especially when we deal with the existence of solution for (0.2), and it seems that rigorous results on the existence of (global) solutions to Boltzmann equation have been obtained only for the cutoff type ([2, 15, 1]).

McKean [9] introduced a class of Markov processes associated with certain nonlinear parabolic equations. The initial value problem for (0.2) is *nearly* the same as the existence problem of the associated Markov process of the type introduced by McKean. In this paper, instead of (0.2) we study its weak version by probabilistic methods:

$$\frac{d}{dt}\langle u,\xi\rangle=\langle u\otimes u,K\xi\rangle,\qquad \xi\in C_0^\infty(\mathbf{R}^3); \tag{0.3}$$

here $C_0^\infty(\mathbf{R}^3)$ is the space of real valued C^∞-functions on R^3 with compact support,

$$(K\xi)(x,x_1)=\int_{(0,\pi)\times(0,2\pi)}\{\xi(x')-\xi(x)\}\,Q(d\theta)\,d\varphi, \tag{0.4}$$

and $\langle u,\xi\rangle$ denotes the integral of ξ with respect to a probability measure solution $u=u(t,\cdot)$ to be sought.

The main objectives of this paper are the followings:

(I) The construction of the Markov process associated with (0.3) by solving certain stochastic differential equation.

(II) The trend to the equilibrium for (0.3).

Chapter I is devoted to the construction of the associated Markov process. A part of the present results was summarized in [18]; here we will give full proofs. The idea is to use the following stochastic differential equation

$$X(t)=X(0)+\int_{(0,t]\times S}a(X(s-),Y(s-,\alpha),\theta,\varphi)\,N(ds\,d\theta\,d\varphi\,d\alpha), \tag{0.5a}$$

or what is the same thing,

$$X(t)=X(0)+\sum_{s\leqq t} a(X(s-),\,Y(s-),\,p(s)); \tag{0.5b}$$

here $S=(0,\pi)\times(0,2\pi)\times(0,1)$ and

(i) $\{p(t), t>0\}$ is a Poisson point process on S with characteristic measure $Q(d\theta)\,d\varphi\,d\alpha$, and $N(ds\,d\theta\,d\varphi\,d\alpha)$ is the corresponding Poisson random measure defined by $N(A)=\sum \mathbb{1}_A(s,p(s))$ [1] for $A\in\mathscr{B}(\mathbf{R}_+\times S)$,

(ii) $\{Y(t,\alpha), t\geqq 0\}$ is a right continuous $\mathbf{R}^3$-valued stochastic process defined on the probability space $\{(0,1), d\alpha\}$ and is equivalent in law to the solution process $\{X(t), t\geqq 0\}$,

(iii) $a(x,x_1,\theta,\varphi)=x'-x$ and $a(X(s-),\ Y(s-),\ \sigma)=a(X(s-), Y(s-,\ \alpha),\ \theta,\ \varphi)$ for $\sigma=(\theta,\varphi,\alpha)\in S$.

Because of the nonlinearity of (0.3), the right hand side of (0.5) involves not only the solution $X(s)$ but also its copy $Y(s)$; in this sense the equation (0.5) is similar to the one considered by McKean [11] in the diffusion case. It will be proved that the equation (0.5) has a solution $\{X(t)\}$ provided the initial distribution has finite expectation, and that the uniqueness in the law sense holds. Also, the solutions to (0.5) will give a Markov process associated with (0.3), the precise definition of which will be given in §1. On the other hand, it was proved in [19] that path functions of any Markov process associated with (0.3) can be represented as solutions to (0.5) after a suitable extension of the basic probability space. Thus we shall obtain the existence and the uniqueness of the associated Markov process.

The existence of the associated Markov process implies that of the associated nonlinear semigroup. Let $\{X(t)\}$ be the associated Markov process (solution of (0.5)) with initial distribution f satisfying $\int|x|f(dx)<\infty$, and denote by $T_t f$ the probability distribution of $X(t)$. Then $u(t)=T_t f$ solves (0.3) and $\{T_t\}$ becomes a nonlinear semigroup. Denote by $\mathscr{P}_2$ the space of probability distributions f on $\mathbf{R}^3$ satisfying $\int|x|^2 f(dx)<\infty$. In Chapter II of this paper, we study $\{T_t\}$ on the space $\mathscr{P}_2$ by making use of the functional $\mathfrak{e}$ and the metric ρ defined on $\mathscr{P}_2$ as follows. For $f_1, f_2\in\mathscr{P}_2$ we put

$$\mathfrak{e}(f_1,f_2)=\inf\int_{\mathbf{R}^3\times\mathbf{R}^3}|x-y|^2\,F(dx\,dy),$$

$$\rho(f_1,f_2)=\sqrt{\mathfrak{e}(f_1,f_2)},$$

where the infimum is taken over all probability measures F in $\mathbf{R}^6$ satisfying $F(A\times\mathbf{R}^3)=f_1(A)$ and $F(\mathbf{R}^3\times A)=f_2(A)$ for any $A\in\mathscr{B}(\mathbf{R}^3)$. For $f\in\mathscr{P}_2$ satisfying $\int|x-m|^2 f(dx)\equiv 3v>0$ where m is the mean vector of f, we put

$$\mathfrak{g}_f(dx)=(2\pi v)^{-3/2}\exp\{-|x-m|^2/2v\}\,dx,$$

$$\mathfrak{e}(f)=\mathfrak{e}(f,\mathfrak{g}_f).$$

It will be seen that ρ gives a metric in $\mathscr{P}_2$. In the one-dimensional case the functional $\mathfrak{e}$ was introduced in connection with the study of Kac's one-dimen-

---

[1] $\mathbb{1}_A$ denotes the indicator function of A throughout

sional model of Maxwellian molecules by Tanaka [17], and a part of the results in [17] (concerning some basic properties of $\mathfrak{e}$ itself) was then extended to the several dimensional case by Murata and Tanaka [12] and to the case of Hilbert spaces by Kondô and Negoro [8].

The main results of Chapter II are as follows:

(A) *The nonlinear semigroup* $\{T_t\}$ *on* $\mathscr{P}_2$ *is non-expansive with respect to the metric* ρ:

$$\rho(T_t f_1, T_t f_2) \leqq \rho(f_1, f_2), \quad t \geqq 0,\ f_1, f_2 \in \mathscr{P}_2.$$

(B) *If* $f \in \mathscr{P}_2$ *satisfies* $\int |x-m|^2 f(dx) \equiv 3v > 0$ *where* m *is the mean vector of* f, *then* $\mathfrak{e}(T_t f)$ *decreases to* 0 *as* $t \uparrow \infty$, *and hence in particular* $T_t f$ *converges to* $\mathfrak{g}_f$ *as* $t \uparrow \infty$.

The (rigorous) entropy arguments in dealing with the trend to equilibrium require the existence of initial densities with finite entropy. According to our method, though it works only for Maxwellian type, we need less restrictions on initial distributions for proving the trend to equilibrium. Also, the result (A) will provide a typical example of a semigroup of nonlinear operators which are non-expansive with respect to certain metric.

I wish to thank T. Ueno; I came to be interested in Maxwellian molecules through conversations with him.

Chapter I. Associated Markov Process

§ 1. Definition of Markov Process Associated with (0.3)

Let us denote by $\mathscr{P}_1$ the family of probability distributions f on $\mathbf{R}^3$ satisfying $\int_{\mathbf{R}^3} |x| f(dx) < \infty$, and introduce the following

Definition. $\{e_f(t, x, \cdot): f \in \mathscr{P}_1, t \geqq 0, x \in \mathbf{R}^3\}$ is called a transition function associated with (0.3), if the following five conditions are satisfied.

(e.1) For fixed $f \in \mathscr{P}_1$, $t \geqq 0$ and $x \in \mathbf{R}^3$, $e_f(t, x, \cdot)$ is a probability measure on $\mathbf{R}^3$.

(e.2) For fixed $A \in \mathscr{B}(\mathbf{R}^3)$, $e_f(t, x, A)$ is jointly measurable in $(f, t, x) \in \mathscr{P}_1 \times \mathbf{R}_+ \times \mathbf{R}^3$, the Borel structure on $\mathscr{P}_1$ being the one induced by the usual vague topology on $\mathscr{P}_1$.

(e.3) For each $t \geqq 0$ and $f \in \mathscr{P}_1$, there exists a constant c depending only upon t and f such that

$$\int_{\mathbf{R}^3} |y|\, e_f(s, x, dy) \leqq c(1 + |x|), \quad 0 \leqq s \leqq t,\ x \in \mathbf{R}^3.$$

(e.4) If we put

$$u(t, \cdot) = \int_{\mathbf{R}^3} f(dx)\, e_f(t, x, \cdot),$$

$$(K_{u(t)} \xi)(x) = \int_{\mathbf{R}^3} (K\xi)(x, x_1)\, u(t, dx_1),$$

then for $\xi \in C_0^\infty(\mathbf{R}^3)$

$$\langle e_f(t,x,\cdot),\xi\rangle = \xi(x) + \int_0^t \langle e_f(s,x,\cdot), K_{u(s)}\xi\rangle\, ds.$$

(e.5) (Kolmogorov-Chapman equation)

$$e_f(t,x,\cdot) = \int_{\mathbf{R}^3} e_f(s,x,dy)\, e_{u(s)}(t-s,y,\cdot), \qquad 0 \leqq s \leqq t,$$

where u is the same as in (e.4).

In the cutoff case $\left(\int_0^\pi Q(d\theta) < \infty\right)$, solutions to (0.3) can easily be obtained from Wild's formula ([21, 10]); a similar formula can also be used to obtain $e_f(t,x,\cdot)$ defined for *all* probability distributions f on $\mathbf{R}^3$. In the non-cutoff case with which we are concerned in this paper, the restriction $f \in \mathscr{P}_1$ is imposed in the above definition since our present method works only under this restriction.

To proceed, let Ω be the space of $\mathbf{R}^3$-valued function on R_+, and denote by $X_t(\omega)$ (or X_t, for short) the value $\omega(t)$ of $\omega(\in\Omega)$ at t. We put $\mathscr{B} = \sigma\{X_t : t < \infty\}$ and $\mathscr{B}_t = \sigma\{X_s : s \leqq t\}$, where $\sigma\{\text{——}\}$ denotes the smallest σ-field on Ω that makes $\{\text{——}\}$ measurable. Now suppose we are given a transition function $\{e_f(t,x,\cdot)\}$ associated with (0.4). Then there exists a unique family $\{P_f, f \in \mathscr{P}_1\}$ of probability measures on $(\Omega, \mathscr{B})$ such that for $A \in \mathscr{B}(\mathbf{R}^3)$

$$P_f\{X_0 \in A\} = f(A),$$
$$P_f\{X_t \in A \mid \mathscr{B}_s\} = e_{u(s)}(t-s, X_s, A), \qquad P_f\text{-a.s.}, \ 0 \leqq s \leqq t.$$

Thus we obtain a (temporally inhomogeneous) Markov process $\{X_t, P_f, f \in \mathscr{P}_1\}$. This is a Markov process which is associated with (0.3).

In the above we have assumed the existence of $\{e_f(t,x,\cdot)\}$, but we do not know its existence in advance; the analytical proof of the existence seems to be difficult. What we are going to do in Chapter I is, as stated in the introduction, to employ the method of stochastic differential equations in order to obtain an associated Markov process.

§ 2. Preliminaries from Poisson Point Process

Suppose we are given a complete probability space $(\Omega, \mathscr{F}, P)$, a Borel subset S of $\mathbf{R}^d$ and an extra point ∂ not belonging to S. An $S \cup \{\partial\}$-valued process $\{p(t,\omega), t > 0\}$ defined on $(\Omega, \mathscr{F}, P)$ is called a *point process* on S, if i) $p(t,\omega)$ is jointly measurable in (t,ω), and ii) the set $\{t : p(t,\omega) \in S\}$ is countable.

Given a point process $\{p(t), t > 0\}$ on S, we put

$$N(A) = \sum_t \mathbb{1}_A(t, p(t)), \qquad A \in \mathscr{B}((0,\infty) \times S)$$

$$N_t(B) = \sum_{0 < s \leqq t} \mathbb{1}_B(p(s)), \quad B \in \mathscr{B}(S), \ t \geqq 0;$$

$N(\cdot)$ is the associated random measure. Let λ be a given σ-finite Borel measure on S. Then, a point process $\{p(t), t>0\}$ is called a *Poisson point process on S with characteristic measure* λ, if for any *disjoint* family $\{A_1, \ldots, A_n\}$ of Borel sets in $(0, \infty) \times S$ such that $\bar{\lambda}(A_k) = \int_{A_k} dt\, d\lambda < \infty$ $(1 \leqq k \leqq n)$ we have

$$P\{N(A_k) = m_k,\ 1 \leqq k \leqq n\} = \prod_{k=1}^{n} \left\{ e^{-\bar{\lambda}(A_k)} \frac{(\bar{\lambda}(A_k))^{m_k}}{m_k!} \right\}$$

for $m_1, \ldots, m_n \in \mathbf{N}$. The following characterization of Poisson point processes is well-known.

Theorem 2.1. *Suppose we are given a point process* $\{p(t), t>0\}$ *on* S, *a* σ*-finite Borel measure* λ *on* S *and an increasing family* $\{\mathscr{F}_t\}$ *of sub-*σ*-fields of* $\mathscr{F}$. *If, for each* $B \in \mathscr{B}(S)$ *with* $\lambda(B) < \infty$, $\{N_t(B) - \lambda(B)t, t \geqq 0\}$ *is an* $\{\mathscr{F}_t\}$*-martingale, then* $\{p(t), t>0\}$ *is a Poisson point process on S with characteristic measure* λ. *In this case,* $\{p(t), t>0\}$ *is said to be* $\{\mathscr{F}_t\}$*-adapted.*

We often deal with integrals of the form

$$\sum_{s \leqq t} A(s, p(s), \omega) = \int_{(0,t] \times S} A(s, \sigma, \omega)\, N(ds\, d\sigma),$$

where $\{p(t), t>0\}$ is a given $\{\mathscr{F}_t\}$-adapted Poisson point process on S with characteristic measure λ and $N(\cdot)$ is the associated Poisson random measure. When $A(t, \sigma, \omega)$ is predictable[2] satisfying the integrability condition $\int_{(0,t] \times S} E|A(s, \sigma, \omega)|\, ds\, \lambda(d\sigma) < \infty$, then

$$E\{\sum_{s \leqq t} A(s, p(s), \omega)\} = E\{\int_{(0,t] \times S} A(s, \sigma, \omega)\, ds\, \lambda(d\sigma)\}, \tag{2.1}$$

and if in addition $A(s, \sigma, \omega)$, $\tau \leqq s \leqq t$, are $\mathscr{F}_\tau$-measurable in ω for some τ, then

$$\begin{aligned} &E\{\exp(i\zeta \sum_{\tau < s \leqq t} A(s, p(s), \omega)) | \mathscr{F}_\tau\} \\ &= \exp\{\int_{(\tau, t] \times S} (e^{i\zeta A(s, \sigma, \omega)} - 1)\, ds\, \lambda(d\sigma)\}, \qquad \zeta \in \mathbf{R}. \end{aligned} \tag{2.2}$$

For $X(t) = X(0) + \sum_{s \leqq t} A(s, p(s), \omega)$ and $\xi \in C^1(\mathbf{R})$ we have

$$\xi(X(t)) = \xi(X(0)) + \sum_{s \leqq t} \{\xi(X(s-) + A(s, p(s), \omega)) - \xi(X(s-))\}. \tag{2.3}$$

The following lemma is also elementary, but we give the proof for completeness.

---

[2] A real valued function $A(t, \sigma, \omega)$ on $\mathbf{R}_+ \times S \times \Omega$ is said to be predictable, if it is measurable with respect to the predictable σ-field; the latter is defined as the smallest σ-field on $\mathbf{R}_+ \times S \times \Omega$ with respect to which all real valued functions $a(t, \sigma, \omega)$ with the following properties (i) and (ii) are measurable.

(i) For each fixed $t \geqq 0$, $a(t, \sigma, \omega)$ is $\mathscr{B}(S) \otimes \mathscr{F}_t$-measurable.

(ii) For each fixed σ and ω, $a(t, \sigma, \omega)$ is left continuous in t.

Lemma 2.1. *Let $A(t,\sigma,\omega)$ be real valued and predictable. Let $T>0$ and assume that there exists $A(\sigma,\omega)$ such that*

$$|A(t,\sigma,\omega)| \leqq A(\sigma,\omega) \quad \text{for } 0 \leqq t \leqq T,$$
$$E\{\int_S A(\sigma,\omega)\,\lambda(d\sigma)\} < \infty. \tag{2.4}$$

Then, for any $\varepsilon>0$ there exists a partition Δ of $[0,T]$:

$$\Delta: 0=t_0<t_1<\cdots<t_n=T$$

such that $|\Delta| = \max(t_k - t_{k-1}) < \varepsilon$ and

$$E\{\int_{(0,T]\times S} |A(t,\sigma,\omega) - A(\Delta(t),\sigma,\omega)|\,dt\,\lambda(d\sigma)\} < \varepsilon,$$

where $\Delta(t)$ is defined by $\Delta(0)=0$ and

$$\Delta(t) = t_{k-1} \quad \text{for } t_{k-1} < t \leqq t_k \ (1 \leqq k \leqq n).$$

Proof. For convenience we redefine $A(t,\sigma,\omega)$ for $t>T$ by putting $A(t,\sigma,\omega)=0$ there, and then extend it to $-\infty<t<\infty$ by putting $A(t,\sigma,\omega)=0$ for $t<0$. For an integer $n \geqq 1$ we put

$$\delta_n(t) = k2^{-n} \quad \text{for } k2^{-n} < t \leqq (k+1)2^{-n} \ (k=0, \pm 1, \dots).$$

Then by (2.4) we have for each t

$$\lim_{n\to\infty} \int_0^1 |A(t+s,\sigma,\omega) - A(\delta_n(t)+s,\sigma,\omega)|\,ds = 0$$

almost everywhere with respect to $\lambda \otimes P$, and hence

$$\lim_{n\to\infty} \int_{(0,1)\times \mathbf{R}\times S\times\Omega} |A(t+s,\sigma,\omega) - A(\delta_n(t)+s,\sigma,\omega)|\,ds\,dt\,\lambda(d\sigma)\,P(d\omega) = 0.$$

Therefore, there exist $s \in (0,1)$ and $n_1 < n_2 < \cdots$ such that

$$\lim_{k\to\infty} \int_{\mathbf{R}\times S\times\Omega} |A(t+s,\sigma,\omega) - A(\delta_{n_k}(t)+s,\sigma,\omega)|\,dt\,\lambda(d\sigma)\,P(d\omega) = 0,$$

or equivalently

$$\lim_{k\to\infty} E\{\int_{\mathbf{R}\times S} |A(t,\sigma,\omega) - A(\delta_{n_k}(t-s)+s,\sigma,\omega)|\,dt\,\lambda(d\sigma)\} = 0.$$

But this formula clearly implies the existence of a partition Δ as stated in the lemma.

§ 3. Two Lemmas

We state two lemmas. The first one is of particular importance.

3.1. We set

$$a(x,x_1,\theta,\varphi) = x' - x,$$

and as a function of φ we extend it to the periodic function on $\mathbf{R}$ with period 2π. This function depends upon the choice of the origin $\varphi=0$ in a spherical coordinate system on the sphere S_{x,x_1}. We can easily see that no choices of the origin $\varphi=0$ imply the smoothness of $a(x,x_1,\theta,\varphi)$ in the variables x and x_1, but we can prove that $a(x,x_1,\theta,\varphi)$ has *a sort of* Lipschitz continuity which is enough for our later developments.

Lemma 3.1. *There exist a constant c and a Borel function* $\varphi_0(x,x_1,y,y_1)$ *on* $\mathbf{R}^{12}$ *such that*

$$|a(x,x_1,\theta,\varphi)-a(y,y_1,\theta,\varphi+\varphi_0(x,x_1,y,y_1))| \leqq c\{|x-y|+|x_1-y_1|\}\,\theta.$$

Proof. (i) When $x=x_1$, we put $\varphi_0(x,x,y,y_1)=0$. Since $a(x,x,\theta,\varphi)=0$, we have

$$\begin{aligned}&|a(x,x,\theta,\varphi)-a(y,y_1,\theta,\varphi+\varphi_0)|\\ &\quad=|a(y,y_1,\theta,\varphi)|\leqq\frac{|y-y_1|}{2}\theta\\ &\quad\leqq\tfrac{1}{2}\{|x-y|+|x_1-y_1|\}\theta.\end{aligned}$$

(ii) When $y=y_1$, we obtain a similar result with $\varphi_0(x,x_1,y,y)=0$.

(iii) We assume that $x\neq x_1$ and $y\neq y_1$. Let l be the straight line which passes through the point $(x+x_1)/2$ and is perpendicular to the plane determined by the three points $(x+x_1)/2$, x and x^*, where

$$x^*=\frac{|x-x_1|}{|y-y_1|}\cdot\frac{y-y_1}{2}+\frac{x+x_1}{2}.$$

We denote by ρ the rotation around l sending x to x^*. Also we define the transformations τ and $\tilde\tau$ from $\mathbf{R}^3$ to itself by

$$\tau z=\frac{|y-y_1|}{|x-x_1|}\left(z-\frac{x+x_1}{2}\right)+\frac{x+x_1}{2},$$

$$\tilde\tau z=z+\frac{y+y_1}{2}-\frac{x+x_1}{2}.$$

Then we have

$$\rho x=x^*,\qquad \tau x^*=x_*\equiv\frac{y-y_1}{2}+\frac{x+x_1}{2},\qquad \tilde\tau x_*=y,$$

and $\tilde\tau\tau\rho$ sends the sphere S_{x,x_1} to the sphere S_{y,y_1}. So, if we put

$$A(x,x_1,\theta,\varphi)=a(x,x_1,\theta,\varphi)+x(=x'),$$

$\tilde\tau\tau\rho A(x,x_1,\theta,0)$ lies on S_{y,y_1} and its longitude is independent of the colatitude θ. Therefore, there exists a function $\varphi_0(x,x_1,y,y_1)$ taking values in $[0,2\pi)$ such that

$$\tilde\tau\tau\rho A(x,x_1,\theta,0)=A(y,y_1,\theta,\varphi_0(x,x_1,y,y_1)).$$

We then have the following formula.

$$\tilde{\tau}\tau\rho A(x,x_1,\theta,\varphi)=A(y,y_1,\theta,\varphi+\varphi_0(x,x_1,y,y_1)).$$

We now claim that

$$|a(x,x_1,\theta,\varphi)-a(y,y_1,\theta,\varphi+\varphi_0(x,x_1,y,y_1))| \leqq(\pi+\tfrac{1}{2})\{|x-y|+|x_1-y_1|\}\,\theta. \tag{3.1}$$

To show this, we first notice that

$$\begin{aligned}|a(x,x_1,\theta,\varphi)-(\rho A(x,x_1,\theta,\varphi)-x^*)| &\leqq |a(x,x_1,\theta,\varphi)|\times(\text{the rotation angle of }\rho)\\ &\leqq\frac{\pi}{2}|x-x^*|\,\theta,\end{aligned} \tag{3.2}$$

$$\begin{aligned}|(\rho A(x,x_1,\theta,\varphi)-x^*)-(\tau\rho A(x,x_1,\theta,\varphi)-\tau x^*)|\\ \leqq\tfrac{1}{2}\{|x-x_1|-|y-y_1|\}\,\theta\leqq\tfrac{1}{2}\{|x-y|+|x_1-y_1\}\,\theta,\end{aligned} \tag{3.3}$$

$$\begin{aligned}\tau\rho A(x,x_1,\theta,\varphi)-\tau x^* &=\tilde{\tau}\tau\rho A(x,x_1,\theta,\varphi)-\tilde{\tau}\tau x^*\\ &=A(y,y_1,\theta,\varphi+\varphi_0(x,x_1,y,y_1))-y\\ &=a(y,y_1,\theta,\varphi+\varphi_0(x,x_1,y,y_1)).\end{aligned} \tag{3.4}$$

From (3.2), (3.3) and (3.4) we then have

$$|a(x,x_1,\theta,\varphi)-a(y,y_1,\theta,\varphi+\varphi_0(x,x_1,y,y_1))| \leqq\frac{\pi}{2}|x-x^*|\,\theta+\tfrac{1}{2}\{|x-y|+|x_1-y_1|\}\,\theta,$$

which combined with the following inequalities proves (3.1):

$$\begin{aligned}|x-x^*| &\leqq |x-y|+|y-x_*|+|x_*-x^*|\\ &\leqq |x-y|+\left|\frac{x+x_1}{2}-\frac{y+y_1}{2}\right|+\left|\frac{|x-x_1|}{|y-y_1|}-1\right|\cdot\frac{|y-y_1|}{2}\\ &\leqq |x-y|+|x-y|+|x_1-y_1|.\end{aligned}$$

3.2. In this paper we often consider stochastic processes having sample paths in the following space W:

$W=$ the space of R^3-valued right continuous functions on $\mathbf{R}_+$ having left limits.

In W we consider the Skorohod topology. Then it is well known that W is a completely metrizable and separable space (see [7, 14]) and that the topological Borel field $\mathscr{B}_W$ on W coincides with the usual coordinate σ-field. We think of the unit interval $(0,1)$ as a probability space by considering the Lebesgue measure (strictly speaking, its restriction to $\mathscr{B}(0,1)$, the σ-field of Borel sets in

$(0,1)$). A stochastic process defined on this probability space and having sample paths in W is called an α-process for simplicity; similarly a random variable on this probability space is called an α-random variable. We sometimes want to have α-processes constructed as in the following way.

Lemma 3.2. *Suppose we are given two processes* $\mathbf{X}_1=\{X_1(t), t\geqq 0\}$ *and* $\mathbf{X}_2=\{X_2(t), t\geqq 0\}$ *defined on a common probability space* $(\Omega, \mathscr{F}, P)$ *and having sample paths in W. Let* $\mathbf{Y}_1=\{Y_1(t,\alpha), t\geqq 0\}$ *be an* α*-process which is equivalent in law to* $\mathbf{X}_1$. *We assume that there exists an* α*-random variable* η *which is independent of* $\mathbf{Y}_1$ *and uniformly distributed on the interval* $(0,1)$. *Then we can construct an* α*-process* $\mathbf{Y}_2=\{Y_2(t,\alpha), t\geqq 0\}$ *in such a way that* (i) *the joint process* $(\mathbf{Y}_1,\mathbf{Y}_2)$ *is equivalent in law to* $(\mathbf{X}_1,\mathbf{X}_2)$ *and* (ii) *there still exists an* α*-random variable which is independent of* $\mathbf{Y}_2$ *and uniformly distributed on* $(0,1)$.

Proof. Denote by U the probability measure on $W\times W$ induced by the joint process $(\mathbf{X}_1,\mathbf{X}_2)$, and by U_1 the one on W induced by $\mathbf{X}_1$. Since W is a complete metric separable space, there exists a transition function $P(w,A)$ of $\mathbf{X}_2$ given $\mathbf{X}_1$ with the following three properties:

For each fixed $w\in W$, $P(w,\cdot)$ is a probability measure on W. (3.5)

For each fixed $A\in\mathscr{B}_W$, $P(\cdot,A)$ is a Borel function on W. (3.6)

For any $A_1, A_2\in\mathscr{B}_W$

$$U(A_1\times A_2)=\int_{A_1} P(w,A_2)\,U_1(dw). \tag{3.7}$$

Since any complete metric separable space having the same cardinality as $\mathbf{R}$ is Borel isomorphic to $\mathbf{R}$ (see [14]), there exists a Borel isomorphism Φ from W into $\mathbf{R}$. We fix such a Φ and put

$$\tilde{Y}(w,\alpha)=\sup\{x: P(w,\Phi^{-1}((-\infty,x]))\leqq\alpha\},$$
$$\mathbf{Y}(w,\alpha)=\Phi^{-1}(\tilde{Y}(w,\alpha)),\qquad \alpha\in(0,1).$$

Then $\mathbf{Y}(w,\alpha)$ is jointly measurable, and for each fixed $w\in W$ the distribution of $Y(w,\cdot)$ on W is $P(w,\cdot)$. Taking two (arbitrary) independent α-random variables η_1 and η_2 with the uniform distribution on $(0,1)$, and regarding $\mathbf{Y}_1(\alpha)=\{Y_1(t,\alpha), t\geqq 0\}$ as an element of W, we can define a W-valued α-random variable $\mathbf{Y}_2$ by $\mathbf{Y}_2(\alpha)=\mathbf{Y}(\mathbf{Y}_1(\alpha), \eta_1(\eta(\alpha)))$. Then the joint process $(\mathbf{Y}_1,\mathbf{Y}_2)$ is clearly U-distributed, and $\eta_2(\eta(\alpha))$ is a uniformly distributed α-random variable independent of $\mathbf{Y}_2$.

§ 4. Stochastic Differential Equation

We use the notations introduced in § 3, such as the function $a(x,x_1,\theta,\varphi)$, the probability space $(0,1)$ and the space W. In this section the probability space $(0,1)$ is of an auxiliary character, and the basic complete probability space is $(\Omega,\mathscr{F},P)$ which is chosen suitably. Let $S=(0,\pi)\times(0,2\pi)\times(0,1)$ and λ be the measure on S defined by $d\lambda = Q(d\theta)\,d\varphi\,d\alpha$, where $Q(d\theta)$ is a given probability measure on $(0, \pi)$ satisfying $\int_0^\pi \theta Q(d\theta) < \infty$.

On a probability space $(\Omega, \mathscr{F}, P)$ suppose we are given (i) an increasing family $\{\mathscr{F}_t\}_{t \geqq 0}$ of sub-σ-fields of $\mathscr{F}$, (ii) an $\{\mathscr{F}_t\}$-adapted Poisson point process $\{p(t), t > 0\}$ on S with characteristic measure λ and (iii) an $\mathscr{F}_0$-measurable $\mathbf{R}^3$-valued random variable X. Let $N(ds\,d\theta\,d\varphi\,d\alpha)$ be the Poisson random measure corresponding to $\{p(t), t > 0\}$. Then X and $\{p(t), t > 0\}$ $(N(ds\,d\theta\,d\varphi\,d\alpha))$ are independent. We now consider the stochastic differential equation

$$dX(t) = a(X(t-), Y(t-), \theta, \varphi)\,dN, \qquad X(0) = X, \tag{4.1}$$

whose precise meaning is

$$X(t) = X + \int_{(0,t] \times S} a(X(s-), Y(s-, \alpha), \theta, \varphi)\,N(ds\,d\theta\,d\varphi\,d\alpha), \qquad \text{a.s.,} \tag{4.2a}$$

or equivalently

$$X(t) = X + \sum_{s \leqq t} a(X(s-), Y(s-), p(s)), \qquad \text{a.s.,} \tag{4.2b}$$

where $\{X(t), t \geqq 0\}$ is to be sought as an $\{\mathscr{F}_t\}$-adapted process with sample paths in W under the condition that $\{Y(t, \alpha), t \geqq 0\}$ is an α-process equivalent in law to $\{X(t), t \geqq 0\}$; the notation $a(x, Y, \sigma)$ for an $\mathbf{R}^3$-valued α-random variable Y is defined by

$$a(x, Y, \sigma) = a(x, Y(\alpha), \theta, \varphi) \qquad \text{for } \sigma = (\theta, \varphi, \alpha) \in S. \tag{4.3}$$

In the right hand sides of (4.2a) and (4.2b) we may (and sometimes do) replace the left limits $X(s-)$ and $Y(s-)$ by $X(s)$ and $Y(s)$, respectively. However, the use of the left limits seems to be suited for the intuitive meaning of the motion: *a particle changes its velocity by the interaction with another similar independent particle.*

We use the following notation. For $f_1, f_2 \in \mathscr{P}_1$ we put

$$\rho_1(f_1, f_2) = \inf_{F \in \mathbf{F}} \int_{\mathbf{R}^3 \times \mathbf{R}^3} |x - y|\,F(dx\,dy), \tag{4.4}$$

where $\mathbf{F} = \mathbf{F}(f_1, f_2)$ is the class of probability measures F on R^6 satisfying $F(A \times \mathbf{R}^3) = f_1(A)$ and $F(\mathbf{R}^3 \times A) = f_2(A)$ for any $A \in \mathscr{B}(\mathbf{R}^3)$. Then it is clear that the infimum in (4.4) is attained at some $F \in \mathbf{F}(f_1, f_2)$. Also, it can be proved that ρ_1 gives a metric in $\mathscr{P}_1$; in fact the triangle inequality can be proved as follows. Given $f_1, f_2, f_3 \in \mathscr{P}_1$, we take $F_1 \in \mathbf{F}(f_1, f_2)$ and $F_2 \in \mathbf{F}(f_2, f_3)$ such that

$$\rho_1(f_1, f_2) = \int |x - y|\,F_1(dx\,dy), \qquad \rho_1(f_2, f_3) = \int |x - y|\,F_2(dx\,dy).$$

We can easily construct a probability measure F on $\mathbf{R}^9$ satisfying $F(\tilde{A} \times \mathbf{R}^3) = F_1(\tilde{A})$ and $F(\mathbf{R}^3 \times \tilde{A}) = F_2(\tilde{A})$ for any $\tilde{A} \in \mathscr{B}(R^6)$, and we have

$$\begin{aligned}\rho_1(f_1, f_2) + \rho_1(f_2, f_3) &= \int |x - y|\,F(dx\,dy\,dz) + \int |y - z|\,F(dx\,dy\,dz)\\ &\geqq \int |x - z|\,F(dx\,dy\,dz) \geqq \rho_1(f_1, f_3).\end{aligned}$$

The existence and uniqueness of the solution to (4.1) are now our objectives. We begin with the uniqueness part. Denote by f the probability distribution of the initial value X.

Lemma 4.1. *Assume that $E\{|X|\}<\infty$, that is, $f\in\mathscr{P}_1$. Let T be any positive constant, Δ a partition of the interval $[0,T]$:*

$$\Delta: 0=t_0<t_1<\cdots<t_n=T, \tag{4.5}$$

and define a process $\{X_\Delta(t), 0\leqq t\leqq T\}$ by

$$\begin{aligned} &X_\Delta(0)=X \\ &X_\Delta(t)=X_\Delta(t_k)+\sum_{t_k<s\leqq t} a(X_\Delta(t_k), Y_k, p(s)) \text{ for } t_k<t\leqq t_{k+1}\ (0\leqq k\leqq n-1), \end{aligned} \tag{4.6}$$

where $Y_0,\dots,Y_{n-1}$ are α-random variables defined in each step so that Y_k has the same probability law as $X_\Delta(t_k)$. Then we have the following assertions.

(i) *The probability law of the process $\{X_\Delta(t), 0\leqq t\leqq T\}$ is uniquely determined by f (and so does not depend upon the choice of $Y_0,\dots,Y_{n-1}$).*

(ii) *Let $X^\#$ be another $\mathscr{F}_0$-measurable random variable with probability distribution $f^\#$ in $\mathscr{P}_1$, and define $\{X^\#(t), 0\leqq t\leqq T\}$ by a rule similar to (4.6) replacing X by $X^\#$. Then, enlarging the probability space if necessary, we can construct two processes $\{X(t), 0\leqq t\leqq T\}$ and $\{\tilde X(t), 0\leqq t\leqq T\}$ which are equivalent in law to $\{X_\Delta(t), 0\leqq t\leqq T\}$ and $\{X_\Delta^\#(t), 0\leqq t\leqq T\}$, respectively, and satisfying*

$$E|X(t)-\tilde X(t)|\leqq e^{c_0 t}\rho_1(f,f^\#) \tag{4.7}$$

with $c_0=4\pi c\int_0^\pi \theta Q(d\theta)$, where c is the constant appearing in Lemma 3.1.

Proof. (i) By (2.2) we have

$$\begin{aligned} &E\{e^{i\zeta\cdot X_\Delta(t)}|\mathscr{F}_{t_k}\} \\ &=\exp\{i\zeta\cdot X_\Delta(t_k)+(t-t_k)\int_S(e^{i\zeta\cdot a(X_\Delta(t_k),y,\theta,\varphi)}-1)u_k(dy)Q(d\theta)d\varphi\} \end{aligned}$$

$$t_k\leqq t\leqq t_{k+1},\ \zeta\in\mathbf{R}^3,$$

where $u_k(dy)$ is the probability distribution of $X_\Delta(t_k)$. Then (i) is clear from this formula.

(ii) Enlarging the probability space if necessary, we can assume that there exists an $f^\#$-distributed random variable $\tilde X$ such that $E|X-\tilde X|=\rho_1(f,f^\#)$. For each k $(0\leqq k<n)$, denote by u_k and $u_k^\#$ the probability distributions of $X_\Delta(t_k)$ and $X_\Delta^\#(t_k)$, respectively, and then take α-random variables Y_k and $\tilde Y_k$ with distributions u_k and $u_k^\#$, respectively, in such a way that $E_\alpha|Y_k-\tilde Y_k|=\rho_1(u_k,u_k^\#)$ holds. We now put $X(0)=X$ and $\tilde X(0)=\tilde X$, and assume that $\{X(t)\}$ and $\{\tilde X(t)\}$ are defined for $0\leqq t\leqq t_k$. We first define $\tilde a(x,\tilde Y_k,\sigma)$ for $\sigma=(\theta,\varphi,\alpha)$ by

$$\begin{aligned} &\tilde a(x,\tilde Y_k,\sigma)=a(x,\tilde Y_k(\alpha),\varphi+\varphi_0), \\ &\qquad\varphi_0=\varphi_0(X(t_k),Y_k(\alpha),\tilde X(t_k),\tilde Y_k(\alpha)), \end{aligned}$$

using the function φ_0 of Lemma 3.1, and then put

$$X(t)=X(t_k)+\sum_{t_k<s\leqq t} a(X(t_k), Y_k, p(s)),$$
$$\tilde{X}(t)=\tilde{X}(t_k)+\sum_{t_k<s\leqq t} \tilde{a}(\tilde{X}(t_k), \tilde{Y}_k, p(s)),$$

for $t_k<t\leqq t_{k+1}$. By virtue of (2.1) and Lemma 3.1 we have for $t_k<t\leqq t_{k+1}$

$$\begin{aligned} E|X(t)-\tilde{X}(t)| &\leqq E|X(t_k)-\tilde{X}(t_k)| \\ &\quad +c_1(t-t_k)\{E|X(t_k)-\tilde{X}(t_k)|+\rho_1(u_k,u_k^{\#})\} \\ &\leqq \{1+2c_1(t-t_k)\}\, E|X(t_k)-\tilde{X}(t_k)|, \end{aligned} \tag{4.8}$$

where $c_1=2\pi c\int_0^\pi \theta Q(d\theta)$. Now (4.7) follows from (4.8). The proof is finished.

Lemma 4.2. *Let T be any positive constant and Δ a partition of the interval $[0, T]$ given by (4.5). Let Y_k, $0\leqq k<n$, be α-random variables with probability distributions u_k in $\mathscr{P}_1$, $0\leqq k<n$. Given an $\mathscr{F}_0$-measurable random variable X with probability distribution f in $\mathscr{P}_1$, we define a process $\{X_\Delta(t), 0\leqq t\leqq T\}$ by*

$$X_\Delta(0)=X$$
$$X_\Delta(t)=X_\Delta(t_k)+\sum_{t_k<s\leqq t} a(X_\Delta(t_k), Y_k, p(s)), \qquad t_k<t\leqq t_{k+1}. \tag{4.9}$$

Then we have the following assertions.

(i) *The probability law of $\{X_\Delta(t), 0\leqq t\leqq T\}$ is uniquely determined by f and u_k, $0\leqq k<n$.*

(ii) *Take another $\mathscr{F}_0$-measurable random variable $X^{\#}$ with probability distribution $f^{\#}$ in $\mathscr{P}_1$ and also α-random variables $Y_k^{\#}$, $0\leqq k<n$, with probability distributions $u_k^{\#}$ in $\mathscr{P}_1$. We define a process $\{X_\Delta^{\#}(t), 0\leqq t\leqq T\}$ by a rule similar to (4.9) making use of $X^{\#}$ and $Y_k^{\#}$. Then, enlarging the probability space if necessary, we can construct two processes $\{\mathring{X}(t), 0\leqq t\leqq T\}$ and $\{\mathring{X}^{\#}(t), 0\leqq t\leqq T\}$ which are equivalent in law to the processes $\{X_\Delta(t), 0\leqq t\leqq T\}$ and $\{X_\Delta^{\#}(t), 0\leqq t\leqq T\}$, respectively, and satisfying*

$$\begin{aligned} &E|\mathring{X}(t)-\mathring{X}^{\#}(t)| \\ &\quad\leqq \{1+c_1(t-t_k)\}\, E|\mathring{X}(t_k)-\mathring{X}^{\#}(t_k)|+c_1(t-t_k)\,\rho_1(u_k,u_k^{\#}), \\ &\qquad t_k<t\leqq t_{k+1} \quad (0\leqq k<n), \end{aligned} \tag{4.10}$$

where $c_1=2\pi c\int_0^\pi \theta Q(d\theta)$. In particular, if $\rho_1(u_k,u_k^{\#})<\varepsilon$ for $0\leqq k<n$, then

$$E|\mathring{X}(t)-\mathring{X}^{\#}(t)|\leqq e^{c_1 t}\{\rho_1(f,f^{\#})+\varepsilon\}. \tag{4.11}$$

The proof of this lemma is similar to that of Lemma 4.1 and so is omitted.

Lemma 4.3. *Let $\{X_\Delta(t), 0\leqq t\leqq T\}$ be the same as in Lemma 4.1. Then, any finite dimensional probability law of $\{X_\Delta(t), 0\leqq t\leqq T\}$ is convergent as $|\Delta|\to 0$. More* *precisely, if* $\Box$: $0=s_0<s_1<\cdots<s_m=T$ *is another partition of* $[0,T]$, *then we can construct two processes* $\{X(t), 0\leqq t\leqq T\}$ *and* $\{\tilde{X}(t), 0\leqq t\leqq T\}$ *which are equivalent in law to* $\{X_\Delta(t), 0\leqq t\leqq T\}$ *and* $\{X_\Box(t), 0\leqq t\leqq T\}$, *respectively, and satisfying*

$$E|X(t)-\tilde{X}(t)|\leqq c_2(|\Delta|+|\Box|), \qquad 0\leqq t\leqq T, \tag{4.12}$$

where

$$c_2=2\pi\int_0^\pi \theta Q(d\theta)\exp\left\{2\pi(1+2c)\,T\int_0^\pi \theta Q(d\theta)\right\}E|X|. \tag{4.13}$$

Proof. We may assume that $\Box$ is a sub-partition of Δ without loss of generality. First we construct $\{X_\Delta(t), 0\leqq t\leqq T\}$ as in (4.6) using auxiliary α-random variables $Y_0,\dots,Y_{n-1}$, and then put $X(t)=X_\Delta(t)$, $0\leqq t\leqq T$. Each Y_k can be arbitrarily chosen under the restriction that it is equivalent in law to $X_\Delta(t_k)$. Now we require, in addition, that each Y_k satisfies the following condition:

There exists an α-random variable which is independent of Y_k and uniformly distibuted on $(0,1)$. (4.14)

The process $\{\tilde{X}(t), 0\leqq t\leqq T\}$ must be constructed more carefully. We put

$$\tilde{X}(t)=X+\sum_{s\leqq t} a(X, Y_0, p(s)), \qquad 0\leqq t\leqq s_1.$$

Assuming that $\tilde{X}(t)$ is defined for $0\leqq t\leqq s_k$, we define $\tilde{X}(t)$ for $s_k<t\leqq s_{k+1}$ as follows. Define k' by $t_{k'}=\max\{t_j: t_j\leqq s_k\}$, and then choose an α-random variable $\tilde{Y}_k$ so that the joint distribution of $(Y_{k'}, \tilde{Y}_k)$ coincides with that of $(X(t_{k'}), \tilde{X}(s_k))$; this is possible by virtue of (4.14). Putting

$$\tilde{a}(\tilde{X}(s_k), \tilde{Y}_k, \sigma)=a(\tilde{X}(s_k), \tilde{Y}_k(\alpha), \theta, \varphi+\varphi_0), \qquad \sigma=(\theta,\varphi,\alpha),$$
$$\varphi_0=\varphi_0(X(t_{k'}), Y_{k'}(\alpha), \tilde{X}(s_k), \tilde{Y}_k(\alpha)),$$

we define $\tilde{X}(t)$ for $s_k<t\leqq s_{k+1}$ by

$$\tilde{X}(t)=\tilde{X}(s_k)+\sum_{s_k<s\leqq t}\tilde{a}(\tilde{X}(s_k), \tilde{Y}_k, p(s)).$$

In this way we can construct $\tilde{X}(t)$ for $0\leqq t\leqq T$, and it is not hard to see that thus constructed $\{\tilde{X}(t), 0\leqq t\leqq T\}$ is equivalent in law to $\{X_\Box(t), 0\leqq t\leqq T\}$.

We assume that $s_k\leqq t\leqq s_{k+1}$ for a moment. Since

$$X(t)=X(s_k)+\sum_{s_k<s\leqq t} a(X(t_{k'}), Y_{k'}, p(s)),$$

using Lemma 3.1 we have

$$\begin{aligned} E|X(t)-\tilde{X}(t)| &\leqq E|X(s_k)-\tilde{X}(s_k)|+E\Big\{\int_{(s_k,t]\times S}|a(X(t_{k'}), Y_{k'},\sigma)-\tilde{a}(\tilde{X}(s_k), \tilde{Y}_k,\sigma)|\,ds\,\lambda(d\sigma)\Big\}\\ &\leqq E|X(s_k)-\tilde{X}(s_k)|+(t-s_k)\,2\pi c\int_0^\pi \theta Q(d\theta)\{E|X(t_{k'})-\tilde{X}(s_k)|+E_\alpha|Y_{k'}-\tilde{Y}_k|\}\\ &=E|X(s_k)-\tilde{X}(s_k)|+c_0(t-s_k)\,E|X(t_{k'})-\tilde{X}(s_k)|, \end{aligned} \tag{4.15}$$

where $c_0 = 4\pi c \int_0^\pi \theta Q(d\theta)$. On the other and

$$E|X(t)| \leqq E|X(t_k)| + (t - t_k) E \int_S |a(X(t_k), Y_k, \sigma)| \lambda(d\sigma)$$
$$\leqq E|X(t_k)| + (t - t_k) E \int_S (|X(t_k) - Y_k|/2) \theta \lambda(d\sigma)$$
$$\leqq E|X(t_k)| + c'(t - t_k) E|X(t_k)|, \qquad t_k < t \leqq t_{k+1},$$

where $c' = 2\pi \int_0^\pi \theta Q(d\theta)$, and hence by Gronwall's inequality

$$E|X(t)| \leqq E|X| e^{c't}, \qquad 0 \leqq t \leqq T.$$

Therefore

$$E|X(s_k) - X(t_{k'})| \leqq c'(s_k - t_{k'}) E|X(t_{k'})|$$
$$\leqq c''(s_k - t_{k'}), \qquad c'' = E|X| c' e^{c'T},$$

and hence

$$E|X(t) - \tilde{X}(t)| + c''|\Delta|$$
$$\leqq \{E|X(s_k) - \tilde{X}(s_k)| + c''|\Delta|\} e^{c_0(t - s_k)}, \qquad s_k < t \leqq s_{k+1},$$

which implies that

$$E|X(t) - \tilde{X}(t)| \leqq c''|\Delta|(e^{c_0 t} - 1), \qquad 0 \leqq t \leqq T, \tag{4.16}$$

as was to be proved.

In what follows, a process $\{X(t)\}$ is said to be *integrable* for simplicity, if $E\{\sup_{0 \leqq s \leqq t} |X(s)|\} < \infty$ for each $t \in \mathbf{R}_+$.

Lemma 4.4. *Given an $\mathscr{F}_0$-measurable random variable X with $E\{|X|\} < \infty$, we assume that there exists an integrable solution $\{X(t), t \geqq 0\}$ of* (4.2). *Let T be any positive constant, Δ a partition of $[0, T]$ and $\{X_\Delta(t), 0 \leqq t \leqq T\}$ a process of* Lemma 4.1. *Then, any finite dimensional probability law of $\{X_\Delta(t), 0 \leqq t \leqq T\}$ converges to the corresponding one of $\{X(t), 0 \leqq t \leqq T\}$ as $|\Delta| \to 0$. More precisely, on a suitable probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$ we can construct two processes $\{\tilde{X}(t), 0 \leqq t \leqq T\}$ and $\{\tilde{X}_\Delta(t), 0 \leqq t \leqq T\}$, which are equivalent in law to $\{X(t), 0 \leqq t \leqq T\}$ and $\{X_\Delta(t), 0 \leqq t \leqq T\}$, respectively, and satisfying*

$$\tilde{E}|\tilde{X}(t) - \tilde{X}_\Delta(t)| \leqq c_2|\Delta|, \qquad 0 \leqq t \leqq T, \tag{4.17}$$

with the constant c_2 given by (4.13).

Proof. Define $\Delta(t)$, $0 \leqq t \leqq T$, by

$$\Delta(0) = 0, \; \Delta(t) = t_k \qquad \text{for } t_k < t \leqq t_{k+1} \; (0 \leqq k < n), \tag{4.18}$$

and put

$$X^{\Delta}(t)=X+\sum_{s\leqq t} a(X(\Delta(s)), Y(\Delta(s)), p(s)).$$

Also, we define $\{X^{\#}(t), 0\leqq t\leqq T\}$ by

$$X^{\#}(0)=X,$$

$$X^{\#}(t)=X^{\#}(t_k)+\sum_{t_k<s\leqq t} \tilde{a}(X^{\#}(t_k), Y(t_k), p(s))$$

$$\text{for } t_k<t\leqq t_{k+1} \ (0\leqq k<n), \tag{4.19}$$

where, in each step, $\tilde{a}(X^{\#}(t_k), Y(t_k), \sigma)$ is defined to be equal to $a(X^{\#}(t_k), Y(t_k, \alpha), \sigma, \varphi+\varphi_0)$ for $\sigma=(\theta, \varphi, \alpha)$ with $\varphi_0=\varphi_0(X(t_k), Y(t_k), X^{\#}(t_k), Y(t_k))$. Then we have for $t_k<t\leqq t_{k+1}$

$$\begin{aligned} &E|X^{\Delta}(t)-X^{\#}(t)| \\ &\leqq E|X^{\Delta}(t_k)-X^{\#}(t_k)|+c_1(t-t_k)E|X(t_k)-X^{\#}(t_k)| \\ &\leqq \{1+c_1(t-t_k)\}E|X^{\Delta}(t_k)-X^{\#}(t_k)|+c_1(t-t_k)E|X(t_k)-X^{\Delta}(t_k)|, \end{aligned} \tag{4.20}$$

where c_1 is the same as in (4.10). Now, if we put

$$\varepsilon(\Delta)=E\int_{(0,T]\times S} |a(X(t), Y(t), \sigma)-a(X(\Delta(t)), Y(\Delta(t)), \sigma)|\,dt\,\lambda(d\sigma),$$

then $E|X(t)-X^{\Delta}(t)|\leqq\varepsilon(\Delta)$ for $0\leqq t\leqq T$, and hence (4.20) yields

$$E|X^{\Delta}(t)-X^{\#}(t)|\leqq\varepsilon(\Delta)(e^{c_1T}-1), \tag{4.21}$$

$$E|X(t)-X^{\#}(t)|\leqq\varepsilon(\Delta)e^{c_1T}. \tag{4.22}$$

Since $\{X^{\#}(t)\}$ is equivalent in law to $\{X^{\#}_{\Delta}(t)\}$ which is defined by a rule similar to (4.19) with $\varphi_0\equiv 0$, we can apply Lemma 4.2 to obtain two processes $\{\mathring{X}(t)\}$ and $\{\mathring{X}^{\#}(t)\}$ which are equivalent in law to $\{X_{\Delta}(t)\}$ and $\{X^{\#}(t)\}$, respectively, and satisfying (4.10). Since u_k and $u_k^{\#}$ are the probability distributions of $\mathring{X}(t_k)$ and $X(t_k)$, the uses of the triangle inequality for ρ_1 and the estimate (4.22) result in

$$\begin{aligned} \rho_1(u_k, u_k^{\#}) &\leqq E|\mathring{X}(t_k)-\mathring{X}^{\#}(t_k)|+E|X^{\#}(t_k)-X(t_k)| \\ &\leqq E|\mathring{X}(t_k)-\mathring{X}^{\#}(t_k)|+\varepsilon(\Delta)e^{c_1T}. \end{aligned}$$

Therefore, (4.10) yields

$$\begin{aligned} E|\mathring{X}(t)-\mathring{X}^{\#}(t)| &\leqq \{1+c_0(t-t_k)\}E|\mathring{X}(t_k)-\mathring{X}^{\#}(t_k)| \\ &\quad +c_1(t-t_k)\varepsilon(\Delta)e^{c_1T}, \qquad t_k<t\leqq t_{k+1}, \end{aligned}$$

which implies that

$$E|\mathring{X}(t)-\mathring{X}^{\#}(t)|\leqq\varepsilon(\Delta)e^{c_1T}(e^{c_0T}-1), \qquad 0\leqq t\leqq T. \tag{4.23}$$

By virtue of (4.22) and (4.23), on a suitable probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$ we can construct two porcesses $\{\tilde{X}(t), 0\leqq t\leqq T\}$ and $\{\tilde{X}_{\Delta}(t), 0\leqq t\leqq T\}$ equivalent in law to

$\{X(t),\ 0\leqq t\leqq T\}$ and $\{X_\Delta(t),\ 0\leqq t\leqq T\}$, respectively, so that they satisfy

$$\begin{aligned}\tilde{E}|\tilde{X}(t)-\tilde{X}_\Delta(t)| &\leqq E|X(t)-X^{\#}(t)|+E|\mathring{X}^{\#}(t)-\mathring{X}(t)|\\ &\leqq \varepsilon(\Delta)\,e^{3c_1T},\qquad 0\leqq t\leqq T.\end{aligned}\tag{4.24}$$

Now by an application of Lemma 2.1 we can make both $\varepsilon(\Delta)$ and $|\Delta|$ arbitrary small, and hence the right hand side of (4.24) tends to 0 as $|\Delta|\to 0$ via some subsequence $\{\Delta_m\}$. Combining this fact with Lemma 4.3, especially with the estimate (4.12), we can easily prove the assertion of the lemma. The proof is finished.

Making use of methods similar to those employed in Lemma 4.3 and 4.4, we obtain the following lemma in which $\{Y(t)\}$ is an α-process given in advance (we do not require that it is equivalent in law to the solution process).

Lemma 4.5. *Given an $\mathscr{F}_0$-measurable random variable X with $E|X|<\infty$ and also an integrable α-process $\{Y(t)\}$ which is continuous in the mean, we assume that there exists an integrable solution $\{X(t)\}$ of*

$$X(t)=X+\sum_{s\leqq t} a(X(s-),\ Y(s-),\ p(s)).$$

Let T be any positive constant, Δ a partition of $[0,T]$ and $\{X_\Delta(t),\ 0\leqq t\leqq T\}$ the process obtained by (4.9) *with $Y_k=Y(t_k)$. Then, any finite dimensional probability law of $\{X_\Delta(t), 0\leqq t\leqq T\}$ converges to the corresponding one of $\{X(t), 0\leqq t\leqq T\}$ as $|\Delta|\to 0$. More precisely, on a suitable probability space $(\tilde{\Omega},\tilde{\mathscr{F}},\tilde{P})$ we can construct two processes $\{\tilde{X}(t),\ 0\leqq t\leqq T\}$ and $\{\tilde{X}_\Delta(t),\ 0\leqq t\leqq T\}$ in such a way that they are equivalent in law to $\{X(t), 0\leqq t\leqq T\}$ and $\{X_\Delta(t), 0\leqq t\leqq T\}$, respectively, and satisfy*

$$\tilde{E}|\tilde{X}(t)-\tilde{X}_\Delta(t)|\leqq c_3|\Delta|+\varepsilon_Y(\Delta)e^{c_1T},\qquad 0\leqq t\leqq T,\tag{4.25}$$

where

$$\begin{aligned}c_3&=\pi\int_0^\pi \theta Q(d\theta)\exp\left\{\pi(1+2c)\,T\int_0^\pi \theta Q(d\theta)\right\}(E|X|+M),\\ M&=\sup_{0\leqq t\leqq T} E_\alpha|Y(t)|,\\ \varepsilon_Y(\Delta)&=\max_{0\leqq k<n} E_\alpha|Y(t_{k+1})-Y(t_k)|.\end{aligned}$$

The following uniqueness theorem follows immediately from Lemma 4.1 and 4.4.

Theorem 4.1. *The uniqueness in the law sense holds for integrable solutions of* (4.1), *that is, the probability law on W of any integrable solution of* (4.1) *is uniquely determined by its initial distribution f if $f\in\mathscr{P}_1$. More precisely, if $\{X(t)\}$ and $\{X^{\#}(t)\}$ are any integrable solutions of* (4.1) *with initial distributions f and $f^{\#}$ $(\in\mathscr{P}_1)$, respectively, then on a suitable probability space $(\tilde{\Omega},\tilde{\mathscr{F}},\tilde{P})$ we can construct two processes $\{\tilde{X}(t)\}$ and $\{\tilde{X}^{\#}(t)\}$ in such a way that they are equivalent in law to $\{X(t)\}$ and $\{X^{\#}(t)\}$, respectively, and satisfy*

$$\tilde{E}|\tilde{X}(t)-\tilde{X}^{\#}(t)|\leqq e^{c_0t}\rho_1(f,f^{\#}),\qquad t\geqq 0.\tag{4.26}$$

Next we deal with the existence theorem concerning (4.1); the precise statement of this is given as follows.

Theorem 4.2. *Let $f \in \mathscr{P}_1$ be given. Then, on a suitable probability space $\{\Omega, \mathscr{F}, P\}$ with an increasing family $\{\mathscr{F}_t\}$ of sub-σ-fields we can construct an $\{\mathscr{F}_t\}$-adapted Poisson point process $\{p(t)\}$ on S with characteristic measure λ so that* (4.1) *has an integrable solution with initial distribution f.*

In order to prove this theorem it is convenient to introduce another stochastic differential equation which will turn out to be essentially the same as (4.1). It is expressed as

$$dX(t) = a(X(t-), Y(t-), \theta, \varphi + \varphi^*)\, dN, \qquad X(0) = X, \tag{4.1*}$$

and a solution $\{X(t)\}$ of this equation should be found as an $\{\mathscr{F}_t\}$-adapted process with sample paths in W under the conditions that $\{Y(t)\}$ is an α-process equivalent in law to $\{X(t)\}$ and that $\varphi^* = \varphi^*(t, \alpha, \omega)$ is an $\{\mathscr{F}_t\}$-predictable process. Always, $\varphi + \varphi^*$ should be interpreted mod 2π. Now, for φ^* appearing in (4.1*) we put

$$N^*(A) = \int_{(0,\infty)\times S} \mathbb{1}_A(t, \theta, \varphi + \varphi^*(t, \alpha, \omega), \alpha)\, N(dt\, d\sigma), \qquad A \in \mathscr{B}((0,\infty)\times S).$$

Then by Theorem 2.1 $\{N^*(dt\, d\sigma)\}$ is again an $\{\mathscr{F}_t\}$-adapted Poisson random measure corresponding to λ, and (4.1*) is nothing but (4.1) with N replaced by N^*. Therefore, for the proof of Theorem 4.2 it is enough to prove the following theorem.

Theorem 4.3. *Let $\{p(t)\}$ be an $\{\mathscr{F}_t\}$-adapted Poisson point process on S with characteristic measure λ, and $\{N(dt\, d\sigma)\}$ the associated Poisson random measure. Then, for any $\mathscr{F}_0$-measurable R^3-valued random variable X with $E|X| < \infty$, there exists an integrable solution of* (4.1*).

Proof. We prove this by iteration. First we take an α-random variable Y with the same distribution as X and also with the property that

$$\text{there exists an } \alpha\text{-random variable which is independent of } Y \text{ and uniformly distributed on } (0.1). \tag{4.27}$$

We then define $\{X_1(t)\}$ by

$$X_1(t) = X + \sum_{s \le t} a_1(X_0(s-), Y_0(s-), p(s)),$$

where $X_0(s) \equiv X$, $Y_0(s) \equiv Y$, $\varphi_0^* \equiv 0$ and $a_1(x, Y, \sigma) = a(x, Y(\alpha), \varphi + \varphi_0^*)\ (= a(x, Y, \sigma))$ for $\sigma = (\theta, \varphi, \alpha)$. Next, assuming that $\{X_k(t)\}$, $1 \le k \le n$, are defined together with auxiliary α-processes $\{Y_k(t)\}$ and $\{\mathscr{F}_t\}$-predictable processes $\{\varphi_k^*(t, \alpha, \omega)\}$, $0 \le k < n$, we choose an α-process $\{Y_n(t)\}$ in such a way that the joint process $\{(Y_{n-1}(t), Y_n(t))\}$ is equivalent in law to $\{(X_{n-1}(t), X_n(t))\}$ and that (4.27) holds with Y replaced by $\{Y_n(t)\}$; this is possible by virtue of Lemma 3.2. Using the function φ_0 of Lemma 3.1, we then put

$$\varphi_n^*(t, \alpha, \omega) = \varphi_{n-1}^*(t, \alpha, \omega) + \varphi_0(X_{n-1}(t-), Y_{n-1}(t-), X_n(t-), Y_n(t-)),$$

$$a_{n+1}(X_n(t-), Y_n(t-), \sigma) = a(X_n(t-), Y_n(t-, \alpha), \theta, \varphi + \varphi_n^*(t, \alpha, \omega)),$$

for $\sigma=(\theta,\varphi,\alpha)$, and define $\{X_{n+1}(t)\}$ by

$$X_{n+1}(t)=X+\sum_{s\leqq t} a_{n+1}(X_n(s-),Y_n(s-),p(s)). \tag{4.28}$$

Thus we obtain a sequence of processes $\{X_n(t)\}$, $n\geqq 1$. By Lemma 3.1 we have

$$\begin{aligned}&|a_{n+1}(X_n(s-),Y_n(s-),\sigma)-a_n(X_{n-1}(s-),Y_{n-1}(s-),\sigma)|\\&\quad\leqq c\{|X_n(s-)-X_{n-1}(s-)|+|Y_n(s-)-Y_{n-1}(s-)|\}\,\theta.\end{aligned} \tag{4.29}$$

and hence

$$\begin{aligned}&E|X_{n+1}(t)-X_n(t)|\\&\quad\leqq cE\int_{(0,t]\times S}\{|X_n(s-)-X_{n-1}(s-)|+|Y_n(s-)-Y_{n-1}(s-)|\}\,\theta\,ds\,\lambda(d\sigma)\\&\quad=c_0\int_0^t E|X_n(s)-X_{n-1}(s)|\,ds.\end{aligned}$$

Since $E|X_1(t)-X|\leqq c_0 E|X|\,t$ we have

$$E|X_{n+1}(t)-X_n(t)|\leqq E|X|(c_0 t)^{n+1}/(n+1)!$$

and hence

$$\begin{aligned}&E\{\sup_{0\leqq s\leqq t}|X_{n+1}(s)-X_n(s)|\}\\&\quad\leqq E\{\sum_{s\leqq t}|a_{n+1}(X_n(s-),Y_n(s-),p(s))-a_n(X_{n-1}(s-),Y_{n-1}(s-),p(s))|\}\\&\quad\leqq E|X|\,(c_0 t)^{n+1}/(n+1)!.\end{aligned}$$

Therefore, $X_n(t)$ converges almost surely to some limit $X(t)$ as $n\to\infty$ uniformly on each finite t-interval, and hence $Y_n(t)$ also converges almost surely to some α-process $\{Y(t)\}$ which is obviously equivalent in law to $\{X(t)\}$. These convergences together with the inequality (4.29) imply the almost sure convergence of $a_n(X_{n-1}(t-),Y_{n-1}(t-),\sigma)$ to some $\{\mathscr{F}_t\}$-predictable limit $a_\infty(t,\sigma,\omega)$. It is then clear that $X(t-)+a_\infty(t,\sigma,\omega)$ lies on the sphere $S_{X(t-),Y(t-)}$ and has the colatitude θ (almost surely). Consequently, there exists an $\{\mathscr{F}_t\}$-predictable process $\varphi^*=\varphi^*(t,\alpha,\omega)$ such that $a_\infty(t,\sigma,\omega)=a(X(t-),Y(t-),\theta,\varphi+\varphi^*)$. Now letting $n\uparrow\infty$ in (4.28), we see that (4.1*) is satisfied by the triple $(X(t),Y(t),\varphi^*)$, or equivalently, that $\{X(t)\}$ is a solution of (4.1*).

As the final task of this section we prove some moment estimates concerning solutions of (4.1).

Theorem 4.4. *Let* $\{X(t)\}$ *be any integrable solution of* (4.1) *with initial value* X *satisfying* $E\{|X|^\nu\}<\infty$ *for some positive integer* ν. *Then we have*

$$E\{|X(t)|^\nu\}\leqq e^{c'_\nu t}E\{|X|^\nu\},\qquad t\geqq 0, \tag{4.30}$$

where $c'_\nu = 3^\nu \nu\pi \int_0^\pi \theta Q(d\theta)$. *When* $\nu = 1, 2$, *we have*

$$E\{X(t)\} = E\{X\}, \qquad t \geqq 0, \tag{4.31}$$

$$E\{|X(t)|^2\} = E\{|X|^2\}, \qquad t \geqq 0. \tag{4.32}$$

Proof. Let $\{X_\Delta(t), 0 \leqq t \leqq T\}$ be the same as in Lemma 4.1. Then, by Lemma 4.4, in order to prove (4.30) it is enough to show

$$E\{|X_\Delta(t)|^\nu\} \leqq e^{c'_\nu t} E\{|X|^\nu\}, \qquad 0 \leqq t \leqq T. \tag{4.33}$$

We first notice that $E\{|X_\Delta(t)|^\nu\}$ is bounded in $t \in [0, T]$; in fact this follows from the fact that the ν-th moment of

$$\sum_{t_k < s \leqq t} |a(X_\Delta(t_k), Y_k, p(s))|, \qquad t_k < t \leqq t_{k+1}$$

conditioned on the σ-field $\mathscr{F}_{t_k}$ is given by

$$\sum_{j=1}^{\nu} \frac{1}{j!} \sum_{\substack{\nu_1 + \cdots + \nu_j = \nu \\ \nu_1, \cdots, \nu_j \geqq 1}} \frac{\nu!}{\nu_1! \ldots \nu_j!} \prod_{l=1}^{j} ((t - t_k) \int_S |a(X_\Delta(t_k), Y_k, \sigma)|^{\nu_l} \lambda(d\sigma)).$$

We next write $a_k = a(X_\Delta(t_k), Y_k, p(s))$, $\tilde{a}_k = a(X_\Delta(t_k), Y_k, \sigma)$ and then apply (2.3) to (4.6); the result is

$$\begin{aligned} |X_\Delta(t)|^\nu &= |X_\Delta(t_k)|^\nu + \sum_{t_k < s \leqq t} \{|X_\Delta(s-) + a_k|^\nu - |X_\Delta(s-)|^\nu\} \\ &\leqq |X_\Delta(t_k)|^\nu + \nu \sum_{t_k < s \leqq t} |a_k| \{|X_\Delta(s-)| + |a_k|\}^{\nu-1}, \qquad t_k < t \leqq t_{k+1}. \end{aligned}$$

Therefore, we have for $t_k < t \leqq t_{k+1}$

$$E\{|X_\Delta(t)|^\nu\} \leqq E\{|X_\Delta(t_k)|^\nu\} + \nu E \int_{(t_k, t] \times S} |\tilde{a}_k| \{|X_\Delta(s)| + |\tilde{a}_k|\}^{\nu-1} ds\, \lambda(d\sigma).$$

On the other hand, since $|\tilde{a}_k| \{|X_\Delta(s)| + |\tilde{a}_k|\}^{\nu-1}$ is dominated by

$$\begin{aligned} &(\theta/2) \{|X_\Delta(t_k)| + |Y_k|\} \{|X_\Delta(s)| + |X_\Delta(t_k)| + |Y_k|\}^{\nu-1} \\ &\quad \leqq 3^{\nu-1} (\theta/2) \{|X_\Delta(s)|^\nu + |X_\Delta(t_k)|^\nu + |Y_k|^\nu\}, \end{aligned}$$

we have for $t_k < t \leqq t_{k+1}$

$$\begin{aligned} &E\{|X_\Delta(t)|^\nu\} \\ &\quad \leqq E\{|X_\Delta(t_k)|^\nu\} + \bar{c}_\nu \int_{(t_k, t] \times (0, 1)} E\{|X_\Delta(s)|^\nu + |X_\Delta(t_k)|^\nu + |Y_k|^\nu\} ds\, d\alpha \\ &\quad = \{1 + 2\bar{c}_\nu (t - t_k)\} E\{|X_\Delta(t_k)|^\nu\} + \bar{c}_\nu \int_{t_k}^{t} E\{|X_\Delta(s)|^\nu\} ds, \\ &\bar{c}_\nu = 3^{\nu-1} \nu\pi \int_0^\pi \theta Q(d\theta). \end{aligned}$$

Now an application of Gronwall's inequality yields (4.33).

When $E\{|X|^2\}<\infty$, we take the expectation in

$$|X(t)|^2=|X|^2+\sum_{s\leqq t}\{|X(s-)+a(X(s-),Y(s-),p(s))|^2-|X(s-)|^2\}$$

to obtain

$$\begin{aligned}
&E\{|X(t)|^2\}\\
&=E\{|X|^2\}+\int_{(0,t]\times S}E\{|X(s)+a(X(s),Y(s),\sigma)|^2-|X(s)|^2\}\,ds\,\lambda(d\sigma)\\
&=E\{|X|^2\}+\int(|x'|^2-|x|^2)\,Q(d\theta)\,d\varphi\,u(s,dx)\,u(s,dx_1)\,ds\\
&=E\{|X|^2\}+\int\frac{|x'|^2+|x_1'|^2-|x|^2-|x_1|^2}{2}\,Q(d\theta)\,d\varphi\,u(s,dx)\,u(s,dx_1)\,ds\\
&=E\{|X|^2\},
\end{aligned}$$

proving (4.32), where $u(s,\cdot)$ denotes the probability distribution of $X(s)$ and the last two inegrals are performed on $(0,\pi)\times(0,2\pi)\times\mathbf{R}^3\times\mathbf{R}^3\times(0,t]$. The equality (4.31) can also be proved by a method similar to the above. The proof is finished.

§ 5. The Transition Function and the Markov Process Associated with (0.3)

In this section we show that the solutions of (4.1) give rise to a Markov process which is associated with (0.3) in the sense of § 1. As in the preceeding section, $\{p(t)\}$ and $\{N(dt\,d\sigma)\}$ stand for an $\{\mathscr{F}_t\}$-adapted Poisson point process on S with characteristic measure λ and the associated Poisson random measure, respectively. By virtue of the uniqueness in the law sense for solutions of (4.1) we may write P_f for the probability distribution on $(W,\mathscr{B}_W)$ induced by any integrable solution of (4.1) with initial distribution $f\in\mathscr{P}_1$. Given $f\in\mathscr{P}_1$, we take a P_f-distributed α-process $\{Y(t)\}$ and consider the stochastic differential equation

$$dX(t)=a(X(t-),Y(t-),\theta,\varphi)\,dN,\qquad X(0)=x. \tag{5.1}$$

Although this equation has the same expression as (4.1) (except for the initial value), it should be noticed that $\{Y(t)\}$ of (5.1) is a given α-process and so, of course, is not required to be equivalent in law to the solution of (5.1). As in the case of (4.1), the stochastic differential equation (5.1) is essentially equivalent to

$$dX(t)=a(X(t-),Y(t-),\theta,\varphi+\varphi^*)\,dN,\qquad X(0)=x, \tag{5.1*}$$

in which $\varphi^*=\varphi^*(t,\alpha,\omega)$ is an $\{\mathscr{F}_t\}$-predictable process.

The existence of a solution of (5.1*) can be proved by a method of iteration similar to that used in solving (4.1*), and from Lemma 4.5 and (i) of Lemma 4.2 it follows that the probability distribution on $(W,\mathscr{B}_W)$ of any integrable solution of (5.1) (or (5.1*)) is uniquely determined by f and x; we denote this probability distribution on $(W,\mathscr{B}_W)$ by P_f^x. Also we denote by $X_t(w)$, or X_t for short, the value $w(t)$ of $w\in W$ at time t, and put $\mathscr{B}_t=\sigma\{X_s: s\leqq t\}$, $\mathscr{B}=\vee\,\mathscr{B}_t\,(=\mathscr{B}_W)$. Here the notation X_t should not be confused with $X(t)$; the former is defined on W while the latter is on Ω. Combining (4.11) with (4.26) and then using Lemma 4.5, we have the following assertion: For any $x, y\in\mathbf{R}^3$ and $f, g\in\mathscr{P}_1$ we can construct two processes $\{\tilde{X}(t)\}$ and $\{\tilde{X}^{\#}(t)\}$ on a suitable probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$ in such a way that their probability distributions on the path space W are P_f^x and P_g^y respectively, and that

$$\tilde{E}|\tilde{X}(t)-\tilde{X}^{\#}(t)|\leqq e^{c_1 t}\{|x-y|+e^{c_0 t}\rho_1(f,g)\}, \quad t\geqq 0.$$

This assertion immediately implies the following lemma, in which we put

$$e_f(t,x,A)=P_f^x\{X_t\in A\}, \quad A\in\mathscr{B}(R^3). \tag{5.2}$$

Lemma 5.1. (i) *For any $B\in\mathscr{B}_W$, $P_f^x(B)$ is jointly measurable in $(f,x)\in\mathscr{P}_1\times\mathbf{R}^3$.*

(ii) *For each $A\in\mathscr{B}(R^3)$, $e_f(t,x,A)$ is jointly measurable in $(f,t,x)\in\mathscr{P}_1\times\mathbf{R}_+\times\mathbf{R}^3$.*

It can be easily verified that the function $e_f(t,x,A)$ of (5.2) satisfies (e.3) of § 1. (e.4) can also be verified by first applying (2.3) to a solution $X(t)$ of (5.1) and then taking the expectation.

Theorem 5.1. *For any $x\in\mathbf{R}^3, f\in\mathscr{P}_1$, $A\in\mathscr{B}(\mathbf{R}^3)$ and $0\leqq t_0<t_1$, we have*

$$P_f^x\{X_{t_1}\in A|\mathscr{B}_{t_0}\}=e_{u(t_0)}(t_1-t_0, X_{t_0}, A), \quad P_f^x\text{-a.s.}, \tag{5.3}$$

$$P_f\{X_{t_1}\in A|\mathscr{B}_{t_0}\}=e_{u(t_0)}(t_1-t_0, X_{t_0}, A), \quad P_f\text{-a.s.}, \tag{5.4}$$

where $u(t_0)=u(t_0,\cdot)=P_f\{X_{t_0}\in\cdot\}$.

Proof. For a fixed $t_0\geqq 0$, if we put

$$p^{\#}(t)=p(t_0+t), \quad t>0,$$
$$\mathscr{F}_0^{\#}=\{\phi,\Omega\}, \quad \mathscr{F}_t^{\#}=\sigma\{p^{\#}(s),\ 0<s\leqq t\},$$

then $\{p^{\#}(t)\}$ is also an $\{\mathscr{F}_t^{\#}\}$-adapted Poisson point process on S with characteristic measure λ. Let $N^{\#}(dt\,d\sigma)$ be the associated Poisson random measure. Taking a P_f-distributed α-process $\{Y(t)\}$, we define a $P_{u(t_0)}$-distributed α-process $\{Y^{\#}(t)\}$ by $Y^{\#}(t)=Y(t_0+t)$, $t\geqq 0$, and then consider the stochastic differential equation

$$dX^{\#}(t)=a(X^{\#}(t-), Y^{\#}(t-), \varphi+\varphi^{\#})\,dN^{\#}, \quad X^{\#}(0)=y, \tag{5.5}$$

where $\varphi^{\#}=\varphi^{\#}(t,\alpha,\omega)$ is a suitable $\{\mathscr{F}_t^{\#}\}$-predictable process. By a method of iteration similar to that used in solving (4.1*), we can construct a family $\{X_y^{\#}(t), y\in\mathbf{R}^3\}$ of integrable solutions of (5.5) together with $\{\mathscr{F}_t^{\#}\}$-predictable processes $\varphi_y^{\#}=\varphi_y^{\#}(t,\alpha,\omega)$, $y\in\mathbf{R}^3$, in such a way that

(i) $X_y^{\#}(t,\omega)$ is $\mathscr{B}(\mathbf{R}^3)\times\mathscr{F}_t^{\#}$-measurable for each fixed $t\geqq 0$,

(ii) $\varphi_y^{\#}(t,\alpha,\omega)$ is $\mathscr{B}(\mathbf{R}^3)\times\mathfrak{F}^{\#}$-measurable where $\mathfrak{F}^{\#}$ is the predictable σ-field on $\mathbf{R}_+\times(0,1)\times\Omega$ corresponding to $\{\mathscr{F}_t^{\#}\}$.

We now take an integrable solution $\{X(t)\}$ of (5.1*) and put

$$X^0(t)=\begin{cases} X(t) & \text{for } 0\leqq t<t_0 \\ X_{X(t_0)}^{\#}(t-t_0) & \text{for } t\geqq t_0. \end{cases}$$

$$\varphi^0(t,\alpha,\omega)=\begin{cases} \varphi^*(t,\alpha,\omega) & \text{for } 0\leqq t<t_0 \\ \varphi_{X(t_0)}^{\#}(t-t_0,\alpha,\omega) & \text{for } t\geqq t_0. \end{cases}$$

Since $\mathscr{F}_{t_0}$ and $\{\mathscr{F}_t^{\#}\}$ are independent and the process $\{X_y^{\#}(t)\}$ is $P_{u(t_0)}^{y}$-distributed, we have for $t_1 > t_0$

$$\begin{aligned}
&P\{X^0(t_1)\in A\,|\,\mathscr{F}_{t_0}\}\\
&=P\{X^{\#}_{X(t_0)}(t_1-t_0)\in A\,|\,\mathscr{F}_{t_0}\}\\
&=P\{X_y^{\#}(t_1-t_0)\in A\}, \quad y=X(t_0), \text{ a.s.}\\
&=P_{u(t_0)}^{X(t_0)}\{X_{t-t_0}\in A\}=e_{u(t_0)}(t-t_0, X(t_0), A), \quad \text{a.s.}
\end{aligned} \tag{5.6}$$

On the other hand, from (i) and (ii) it follows that $\{X^0(t)\}$ is $\{\mathscr{F}_t\}$-adapted, φ^0 is $\{\mathscr{F}_t\}$-predictable and that $\{X^0(t)\}$ is a solution of (5.1*) with $\varphi^* = \varphi^0$. Thus $\{X^0(t)\}$ is P_f^x-distributed and hence (5.6) implies (5.3). (5.4) can also be proved in a similar manner. The proof is finished.

From what we have proved, it is now clear that $e_f(t, x, A)$ is a transition function associated with (0.3) and so the corresponding Markov process $\{X_t, P_f, f\in\mathscr{P}_1\}$ is also associated with (0.3). As stated in the introduction, it was proved in [19] that this Markov process is the unique one which is associated with (0.3) in the sense of § 1.

Chapter II. Trend to Equilibrium

§ 6. Some Lemmas Concerning ρ-metric on $\mathscr{P}_2$

The purpose of this section is to prepare some lemmas concerning e and ρ, defined below, for the use in later sections. We denote by $\mathscr{P}_2$ the space of probability distributions on $\mathbf{R}^3$ with finite second moments. For f and g in $\mathscr{P}_2$ we put

$$\mathfrak{E}(F) = \int_{R^3\times R^3} |x-y|^2\, F(dx\,dy),$$

$$\mathrm{e}(f,g) = \inf_{F\in\mathbf{F}(f,g)} \mathfrak{E}(F), \qquad \rho(f,g) = \sqrt{\mathrm{e}(f,g)},$$

where $\mathbf{F}(f,g)$ denotes the family of probability distributions F on $\mathbf{R}^6$ satisfying $F(A \times \mathbf{R}^3) = f(A)$ and $F(\mathbf{R}^3 \times A) = g(A)$ for any $A\in\mathscr{B}(R^3)$. Since $\mathbf{F}(f,g)$ is compact with respect to the topology induced by the usual convergence as probability distributions on $\mathbf{R}^6$ and since $\mathfrak{E}(F)$ is continuous on $\mathbf{F}(f,g)$, the infimum value $\mathrm{e}(f,g)$ is attained at some $F\in\mathbf{F}(f,g)$. As in the case of ρ_1 of §4, it can be proved that ρ gives a metric on $\mathscr{P}_2$. However, when we speak of a convergence in $\mathscr{P}_2$, we always mean that it is the usual one as probability distributions unless ρ-convergence is explicitly stated.

The proof of the following lemma is elementary and so is omitted.

Lemma 6.1. (i) *The ρ-convergence implies the usual convergence in $\mathscr{P}_2$.*

(ii) *If $f_n \to f$ in $\mathscr{P}_2$ and if*

$$\lim_{N\to\infty} \sup_{n\geq 1} \int_{|x|>N} |x|^2 f_n(dx) = 0, \tag{6.1}$$

then $\{f_n\}$ is also ρ-convergent to f. In general, $f_n \to f$ and $g_n \to g$ in $\mathscr{P}_2$ imply $\mathrm{e}(f,g) \leq \underline{\lim}_{n\to\infty}\, \mathrm{e}(f_n, g_n)$.

Let $\varepsilon>0$ be fixed. We denote by $\mathfrak{g}_\varepsilon$ the Gaussian distribution $(2\pi\varepsilon)^{-3/2}\exp(-|x|^2 2\varepsilon)\,dx$ on $\mathbf{R}^3$ and put

$$\mathscr{P}_{(\varepsilon)}=\{f*\mathfrak{g}_\varepsilon | f\in\mathscr{P}_2 \text{ and } f(\{|x|>1/\varepsilon\})=0\}.$$

Then we have the following lemma.

Lemma 6.2. *For each pair $(f,g)\in\mathscr{P}_{(\varepsilon)}\times\mathscr{P}_{(\varepsilon)}$, there exists a unique $F_{f,g}\in\mathbf{F}(f,g)$ such that $\mathfrak{E}(F_{f,g})=\mathfrak{e}(f,g)$. Moreover, the mapping Φ from $\mathscr{P}_{(\varepsilon)}\times\mathscr{P}_{(\varepsilon)}$ into the space $\mathscr{P}(\mathbf{R}^6)$ of probability distributions on $\mathbf{R}^6$, defined by $\Phi(f,g)=F_{f,g}$, is continuous.*

Proof. First we remark that

if $F\in\mathbf{F}(f,g)$ and $\mathfrak{E}(F)=\mathfrak{e}(f,g)$, then

$F(A\times B)=\int_B \delta_{\psi(x)}(A)\,g(dx)$ for any $A,B\in\mathscr{B}(R^3)$ with a

suitable Borel mapping ψ from $\mathbf{R}^3$ into itself, (6.2)

where $\delta_{\psi(x)}(\cdot)$ denotes the δ-distribution at $\psi(x)$. In fact, (6.2) was proved in [12: Theorem 1] in the special case when g is the Gaussian distribution with the same mean vector and variance matrix as those of f, and the proof in [12] is also adapted, without any change, to the more general case when g has a strictly positive density with respect to the Lebesgue measure. Next, assume that F_1 and F_2 are in $\mathbf{F}(f,g)$ and satisfy $\mathfrak{E}(F_1)=\mathfrak{E}(F_2)=\mathfrak{e}(f,g)$. Then, $F=(F_1+F_2)/2$ also belongs to $\mathbf{F}(f,g)$ and satisfies $\mathfrak{E}(F)=\mathfrak{e}(f,g)$, and so by (6.2)

$$\delta_{\psi(x)}(A)=\{\delta_{\psi_1(x)}(A)+\delta_{\psi_2(x)}(A)\}/2, \qquad g\text{-a.s.}$$

with some Borel mappings ψ, ψ_1 and ψ_2. But this formula clearly implies that $\psi=\psi_1=\psi_2$, g-a.s., and hence $F_1=F_2$. This proves the first half of the lemma. Finally, to prove the second half, we assume that $f_n\to f$ and $g_n\to g$ in $\mathscr{P}_{(\varepsilon)}$, and write $F_n=F_{f_n,g_n}$. Obviously $\{F_n\}$ is relatively compact in $\mathscr{P}(\mathbf{R}^6)$. Let F be any limit point of $\{F_n\}$. Then by (ii) of Lemma 6.1 we have

$$\mathfrak{e}(f,g)=\lim_{n\to\infty}\mathfrak{e}(f_n,g_n)=\lim_{n\to\infty}\mathfrak{E}(F_n)=\mathfrak{E}(F),$$

which implies that $F=F_{f,g}$ by the uniqueness part of the lemma, proving the continuity of Φ. Thus the proof of the lemma is finished.

Lemma 6.3. *Let $(\Omega,\mathscr{F},P)$ be an arbitrary probability space and suppose that we are given sub-families $\{f^\omega,\omega\in\Omega\}$ and $\{g^\omega,\omega\in\Omega\}$ of $\mathscr{P}_2$ satisfying the following conditions.*

(i) *For each $A\in\mathscr{B}(R^3)$, $f^\omega(A)$ and $g^\omega(A)$ are $\mathscr{F}$-measurable in ω.*

(ii) *The probability distributions $f=\int_\Omega f^\omega P(d\omega)$ and $g=\int_\Omega g^\omega P(d\omega)$ belong to $\mathscr{P}_2$.*

Then we have

$$\mathfrak{e}(f,g)\leqq E\{\mathfrak{e}(f^\omega,g^\omega)\}. \tag{6.3}$$

Proof. For each $\varepsilon>0$ and $f\in\mathscr{P}_2$ let f_ε stand for the probability distribution $\tilde{f}*g_\varepsilon$, where $*$ denotes convolution and

$$\tilde{f}(A)=f(A\cap\{|x|\leqq 1/\varepsilon\})+f(\{|x|>1/\varepsilon\})\,\delta_0(A), \qquad A\in\mathscr{B}(\mathbf{R}^3).$$

Then we have $f_\varepsilon\to f$ as $\varepsilon\downarrow 0$, $\lim\limits_{N\to\infty}\sup\limits_{0<\varepsilon<1}\int\limits_{|x|>N}|x|^2 f_\varepsilon(dx)=0$, and hence $\mathfrak{e}(f_\varepsilon, g_\varepsilon)\to \mathfrak{e}(f,g)$ as $\varepsilon\downarrow 0$ for any $f,g\in\mathscr{P}_2$ by (ii) of Lemma 6.1. Next, denote by F_ε^ω the unique probability distribution on $\mathbf{R}^6$ such that $F_\varepsilon^\omega\in\mathbf{F}(f_\varepsilon^\omega, g_\varepsilon^\omega)$ and $\mathfrak{E}(F_\varepsilon^\omega)=\mathfrak{e}(f_\varepsilon^\omega, g_\varepsilon^\omega)$. Since each mapping in

$$\omega\in\Omega\to(f^\omega,g^\omega)\in\mathscr{P}_2\times\mathscr{P}_2\to(f_\varepsilon^\omega,g_\varepsilon^\omega)\in\mathscr{P}_{(\varepsilon)}\times\mathscr{P}_{(\varepsilon)}\to F_\varepsilon^\omega\in\mathscr{P}(\mathbf{R}^6)$$

is measurable (the last mapping is continuous and hence Borel measurable according to the preceeding lemma), F_ε^ω is also measurable in ω. Therefore $\mathfrak{E}(F_\varepsilon^\omega)\equiv \mathfrak{e}(f_\varepsilon^\omega, g_\varepsilon^\omega)$ is $\mathscr{F}$-measurable in ω, and it follows that

$$\lim_{\varepsilon\downarrow 0}E\{\mathfrak{e}(f_\varepsilon^\omega,g_\varepsilon^\omega)\}=E\{\mathfrak{e}(f^\omega,g^\omega)\}, \tag{6.4}$$

because the integrand $\mathfrak{e}(f_\varepsilon^\omega,g_\varepsilon^\omega)$ is dominated by $2\int|x|^2 f^\omega(dx)+2\int|x|^2 g^\omega(dx)$ which is P-integrable by the assumption (ii) of the lemma. On the other hand, $F_\varepsilon=\int\limits_\Omega F_\varepsilon^\omega\,P(d\omega)$ clearly belongs to $\mathbf{F}(f_\varepsilon, g_\varepsilon)$ and hence we have

$$\mathfrak{e}(f_\varepsilon,g_\varepsilon)\leqq\mathfrak{E}(F_\varepsilon)=\int\limits_\Omega\mathfrak{E}(F_\varepsilon^\omega)\,P(d\omega)=E\{\mathfrak{e}(f_\varepsilon^\omega,g_\varepsilon^\omega)\}.$$

Now letting $\varepsilon\downarrow 0$ in the above and then noting (6.4), we obtain (6.3).

To state the last lemma of this section, let us denote by $C_{x,r,l}$ the circle with center $x\,(\in\mathbf{R}^3)$, of radius r and lying on a plane which is perpendicular to a unit vector l. Also we denote by $U_{x,r,l}$ the uniform distribution on $C_{x,r,l}$; this can be regarded as a probability distribution on $\mathbf{R}^3$ and so $U_{x,r,l}\in\mathscr{P}_2$.

Lemma 6.4. *For any $x,y\in\mathbf{R}^3$, $r,s>0$ and unit vectors l and m, we have*

$$\mathfrak{e}(U_{x,r,l},U_{y,s,m})\leqq|x-y|^2+r^2+s^2-rs\{1+|(l,m)|\}.$$

Proof. In proving the lemma, without loss of generality we may assume that $x=0$, $l=(0,0,1)$ and $m=(0,-\sin\gamma,\cos\gamma)$ with $0\leqq\gamma\leqq\pi/2$. Let $\Omega=[0,2\pi)$, P be the Lebesgue measure in $[0,2\pi)$ multiplied by $1/2\pi$ and put

$$X(\omega)=(r\cos\omega,\ r\sin\omega,\ 0),$$

$$Y(\omega)=y+(s\cos\omega,\ s\sin\omega\cos\gamma,\ s\sin\omega\sin\gamma), \qquad \omega\in\Omega.$$

Then X and Y are random variables that are uniformly distributed on $C_{x,r,l}$ and $C_{y,s,m}$, respectively. Therefore,

$$\mathfrak{e}(U_{x,r,l},U_{y,s,m})\leqq E\{|X-Y|^2\}=\frac{1}{2\pi}\int\limits_0^{2\pi}|X(\omega)-Y(\omega)|^2\,d\omega.$$

By elementary calculations we see that the last term is equal to

$$|y|^2+r^2+s^2-rs(1+\cos\gamma),$$

completing the proof.

§ 7. Non-expansive Property of the Associated Nonlinear Semigroup with Respect to the Metric ρ

Let $\mathbf{X}=\{X_t, P_f, f\in\mathscr{P}_1\}$ be the Markov process associated with (0.3). We associate with each $t\geqq 0$ and $f\in\mathscr{P}_1$ the probability distribution $T_t f$ on $\mathbf{R}^3$ defined by

$$(T_t f)(A)=P_f\{X_t\in A\}, \qquad A\in\mathscr{B}(\mathbf{R}^3).$$

Then, the Markovian property of $\mathbf{X}$ implies the semigroup property of $\{T_t\}$, that is,

$$T_{t+s}f=T_t\,T_s f, \qquad t,s\geqq 0,\ f\in\mathscr{P}_1.$$

Since $f\in\mathscr{P}_2$ implies $T_t f\in\mathscr{P}_2$ by Theorem 4.4, $\{T_t\}$ is also a nonlinear semigroup on $\mathscr{P}_2$. The purpose of this section is to prove that T_t is non-expansive with respect to the metric ρ on $\mathscr{P}_2$.

First we prepare a lemma of an approximation type. Namely, we prove that the Markov process associated with (0.3) can be approximated in an appropriate sense by the one associated with

$$\frac{d}{dt}\langle u,\xi\rangle=\langle u\otimes u, K_\varepsilon\xi\rangle, \qquad \xi\in C_0^\infty(\mathbf{R}^3) \tag{7.1}$$

for small $\varepsilon>0$, where $K_\varepsilon\xi$ is defined by (0.4) with the replacement of $Q(d\theta)$ by $Q_\varepsilon(d\theta)\equiv\mathbb{1}_{(\varepsilon,\pi)}(\theta)Q(d\theta)$. As in § 4, we take an $\{\mathscr{F}_t\}$-adapted Poisson random measure $\{N(dt\,d\sigma)\}$ corresponding to the measure λ. Then the Markov process associated with (7.1) can be obtained by the family of solutions of

$$dX(t)=a_\varepsilon(X(t-), Y(t-), \theta, \varphi)\,dN, \tag{7.2}$$

where $a_\varepsilon(x,x_1,\theta,\varphi)=\mathbb{1}_{(\varepsilon,\pi)}(\theta)\,a(x,x_1,\theta,\varphi)$ and $\{Y(t)\}$ is an α-process which is required to be equivalent in law to the $\{\mathscr{F}_t\}$-adapted solution $\{X(t)\}$ as in the case of (4.1). We denote by $T_t^{(\varepsilon)}f$ the probability distribution of the solution, at time t, of (7.2) with initial distribution $f\in\mathscr{P}_1$.

The proof of the following lemma is slightly complicated, but it can be done in a manner similar to that of Lemma 4.1 and so is omitted.

Lemma 7.1. *Let T be any positive number, Δ a partition of the interval $[0,T]$ given by (4.5) and define a process $\{X_\Delta(t), 0\leqq t\leqq T\}$ by (4.6). Also for a given $\varepsilon\in(0,\pi)$ define a process $\{X_\Delta^\varepsilon(t),\ 0\leqq t\leqq T\}$ by*

$$X_\Delta^\varepsilon(0)=X$$

$$X_\Delta^\varepsilon(t)=X_\Delta^\varepsilon(t_k)+\sum_{t_k<s\leqq t}a_\varepsilon(X_\Delta^\varepsilon(t_k), Y_k^\varepsilon, p(s)) \quad \text{for } t_k<t<t_{k+1}\ \ (0\leqq k<n),$$

where $Y_0^\varepsilon, \dots, Y_{n-1}^\varepsilon$ *are* α*-random variables defined in each step so that* Y_k^ε *has the same probability law as* $X_\Delta^\varepsilon(t_k)$. *Then, if* $E\{|X|^2\} < \infty$, *we can construct two processes* $\{\tilde{X}_\Delta(t), 0 \leqq t \leqq T\}$ *and* $\{\tilde{X}_\Delta^\varepsilon(t), 0 \leqq t \leqq T\}$ *in such a way that they are equivalent in law to* $\{X_\Delta(t), 0 \leqq t \leqq T\}$ *and* $\{X_\Delta^\varepsilon(t), 0 \leqq t \leqq T\}$ *respectively and satisfy*

$$E\{|\tilde{X}_\Delta(t) - \tilde{X}_\Delta^\varepsilon(t)|^2\} \leqq const \int_0^\varepsilon \theta Q(d\theta).$$

Here, const depends on T but neither on ε *nor on* Δ.

The following approximation lemma is an immediate consequence of the above and Lemma 4.4

Lemma 7.2. (i) *Let T be a positive number,* $\varepsilon \in (0, \pi)$ *and* $f \in \mathscr{P}_2$. *Then, on a suitable probability space* $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$ *we can construct two processes* $\{\tilde{X}(t), 0 \leqq t \leqq T\}$ *and* $\{\tilde{X}^\varepsilon(t), 0 \leqq t \leqq T\}$ *in such a way that they are equivalent in law to solutions of* (4.1) *and* (7.2), *respectively, with initial distribution f and satisfy*

$$\tilde{E}\{|\tilde{X}(t) - \tilde{X}^\varepsilon(t)|^2\} \leqq const \int_0^\varepsilon \theta Q(d\theta)$$

with const depending on T but not on ε.

(ii) $\rho(T_t f, T_t^{(\varepsilon)} f) \to 0$ *as* $\varepsilon \downarrow 0$ *for each* $t \geqq 0$ *and* $f \in \mathscr{P}_2$.

Before stating the theorem of this section, we introduce some notations. For each $\theta \in (0, \pi)$ and $x, x_1 \in \mathbf{R}^3$, we put

$$\Pi_{x, x_1, \theta} = U_{z, r, l} = \text{the uniform distribution on } C_{z, r, l},$$

where $z = \{x + x_1 + (x - x_1)\cos\theta\}/2$, $r = |x - x_1|(\sin\theta)/2$ and $l = (x - x_1)/|x - x_1|$, and regard $\Pi_{x, x_1, \theta}$ as a probability distribution on $\mathbf{R}^3$. For any probability distributions f, g on $\mathbf{R}^3$ and $\theta \in (0, \pi)$, we define another probability distribution $(f \circ g)_\theta$ on $\mathbf{R}^3$ by

$$(f \circ g)_\theta(A) = \int_{\mathbf{R}^3 \times \mathbf{R}^3} \Pi_{x, x_1, \theta}(A) f(dx) g(dx_1), \qquad A \in \mathscr{B}(\mathbf{R}^3).$$

Obviously

$$\langle (f \circ g)_\theta, \xi \rangle = \int_{(0, 2\pi) \times \mathbf{R}^3 \times \mathbf{R}^3} \xi(x') d\varphi f(dx) g(dx_1), \qquad \xi \in C_0^\infty(\mathbf{R}^3),$$

and hence $(f \circ g)_\theta \in \mathscr{P}_2$ provided $f, g \in \mathscr{P}_2$. We write $[f]_\theta = (f \circ f)_\theta$ for short, and put

$$\bar{\mathfrak{e}}_\theta(f, g) = \mathfrak{e}(f, g) - \mathfrak{e}([f]_\theta, [g]_\theta).$$

Theorem 7.1. *For each* $t \geqq 0$ T_t *is non-expansive on* $\mathscr{P}_2$ *with respect to the metric* ρ, *that is,*

$$\rho(T_t f, T_t g) \leqq \rho(f, g), \qquad f, g \in \mathscr{P}_2.$$

More precisely, for any $f, g \in \mathscr{P}_2$ we have

$$\bar{e}_\theta(f,g) \geqq 0, \tag{7.3}$$

$$e(T_t f, T_t g) \leqq e(f,g) - 2\pi \int_0^t ds \int_0^\pi \bar{e}_\theta(T_s f, T_s g) Q(d\theta). \tag{7.4}$$

Our discussions are devided into two cases according whether

$$\int_0^\pi Q(d\theta) < \infty \qquad \text{or } = \infty.$$

Case I. First we discuss the special case in which $q \equiv 2\pi \int_0^\pi Q(d\theta) < \infty$. In this case, for each pair of probability distributions f and g on $\mathbf{R}^3$ we can define a probability distribution $f \circ g$ on $\mathbf{R}^3$ by

$$f \circ g = (2\pi/q) \int_0^\pi (f \circ g)_\theta Q(d\theta).$$

With this notation the equation (0.3) is equivalent to

$$\frac{d}{dt}\langle u, \xi \rangle = \langle q(u \circ u - u), \xi \rangle, \qquad \xi \in C_0^\infty(\mathbf{R}^3). \tag{7.5}$$

A unique (probability) solution $u(t)$ of (7.5) for any given initial distribution f can be obtained by a method of iteration, and the solution is explicitly expressed by the so-called Wild sum ([21]):

$$u(t) = e^{-qt} \sum_{n=1}^\infty (1 - e^{-qt})^{n-1} f^{(n)}.$$

Here $f^{(n)}$, $n \geqq 1$, are probability measures on $\mathbf{R}^3$ defined inductively by

$$f^{(1)} = f,$$

$$f^{(n)} = \frac{1}{n-1} \sum_{k=1}^{n-1} f^{(k)} \circ f^{(n-k)}, \qquad n > 1.$$

On the other hand, from what we have proved in Chapter I we know that $T_t f$ is also a solution of (7.5) with initial distribution f, at least if $f \in \mathscr{P}_1$. Therefore, we have

$$T_t f = e^{-qt} \sum_{n=1}^\infty (1 - e^{-qt})^{n-1} f^{(n)}, \qquad f \in \mathscr{P}_1.$$

The proof of the theorem in Case I will be based on the above Wild sum and the following three lemmas.

Lemma 7.3.

$$e(\Pi_{x, x_1, \theta}, \Pi_{y, y_1, \theta}) \leqq \Phi_\theta(x, x_1, y, y_1),$$

where

$$\Phi_\theta(x,x_1,y,y_1)=\left|\frac{1+\cos\theta}{2}(x-y)+\frac{1-\cos\theta}{2}(x_1-y_1)\right|^2$$
$$+\frac{\sin^2\theta}{4}\{|x-x_1|^2+|y-y_1|^2-|x-x_1||y-y_1|-|(x-x_1,y-y_1)|\}.$$

Proof. It is enough to apply Lemma 6.4 with the replacements:

$$x\to\{x+x_1+(x-x_1)\cos\theta\}/2,\qquad y\to\{y+y_1+(y-y_1)\cos\theta\}/2,$$
$$r\to|x-x_1|(\sin\theta)/2,\qquad s\to|y-y_1|(\sin\theta)/2,$$
$$l\to(x-x_1)/|x-x_1|,\qquad m\to(y-y_1)/|y-y_1|.$$

Lemma 7.4. *Let f_1, f_2, g_1 and g_2 belong to $\mathscr{P}_2$. Then we have the following inequalities.*

(i) $\mathrm{e}[(f_1\circ f_2)_\theta,(g_1\circ g_2)_\theta]\leqq\frac{1+\cos\theta}{2}\mathrm{e}(f_1,g_1)+\frac{1-\cos\theta}{2}\mathrm{e}(f_2,g_2),\qquad \theta\in(0,\pi).$

(ii) $\mathrm{e}(f_1\circ f_2,g_1\circ g_2)\leqq\gamma\,\mathrm{e}(f_1,g_1)+(1-\gamma)\,\mathrm{e}(f_2,g_2),$ *where*
$\gamma=(2\pi/q)\int_0^\pi 2^{-1}(1+\cos\theta)\,Q(d\theta).$

Proof. We choose two pairs $\{X_1,Y_1\}$ and $\{X_2,Y_2\}$ of random variables so that they satisfy the following three conditions.

(a) $E\{|X_1-Y_1|^2\}=\mathrm{e}(f_1,g_1),\quad E\{|X_2-Y_2|^2\}=\mathrm{e}(f_2,g_2).$

(b) For $i=1,2$, X_i is f_i-distributed while Y_i is g_i-distributed.

(c) $\{X_1,Y_1\}$ and $\{X_2,Y_2\}$ are independent.

Then we have

$$(f_1\circ f_2)_\theta=E\{\Pi_{X_1,X_2,\theta}\},\qquad (g_1\circ g_2)_\theta=E\{\Pi_{Y_1,Y_2,\theta}\},$$

and hence by Lemma 6.3 and 7.3

$$\mathrm{e}[(f_1\circ f_2)_\theta,(g_1\circ g_2)_\theta]$$
$$\leqq E\{\mathrm{e}(\Pi_{X_1,X_2,\theta},\Pi_{Y_1,Y_2,\theta})\}\leqq E\{\Phi_\theta(X_1,X_2,Y_1,Y_2)\}$$
$$=E\left\{\left|\frac{1+\cos\theta}{2}(X_1-Y_1)+\frac{1-\cos\theta}{2}(X_2-Y_2)\right|^2\right\}$$
$$+\frac{\sin^2\theta}{4}E\{|X_1-X_2|^2+|Y_1-Y_2|^2-|X_1-X_2||Y_1-Y_2|$$
$$-|(X_1-X_2,Y_1-Y_2)|\}.$$

We now use the inequality

$$(X_1-X_2,Y_1-Y_2)\leqq|X_1-X_2|\,|Y_1-Y_2| \tag{7.6}$$

to obtain

$$\begin{aligned}
&\mathrm{e}[(f_1\circ f_2)_\theta,(g_1\circ g_2)_\theta]\\
&\leqq E\left\{\left|\frac{1+\cos\theta}{2}(X_1-Y_1)+\frac{1-\cos\theta}{2}(X_2-Y_2)\right|^2\right\}\\
&\quad+\frac{\sin^2\theta}{4}E\{|X_1-X_2|^2+|Y_1-Y_2|^2-2(X_1-X_2,Y_1-Y_2)\}\\
&=\frac{(1+\cos\theta)^2}{4}E\{|X_1-Y_1|^2\}+\frac{(1-\cos\theta)^2}{4}E\{|X_2-Y_2|^2\}\\
&\quad+\frac{\sin^2\theta}{4}E\{|X_1-Y_1|^2\}+\frac{\sin^2\theta}{4}E\{|X_2-Y_2|^2\}\\
&=\frac{1+\cos\theta}{2}\mathrm{e}(f_1,g_1)+\frac{1-\cos\theta}{2}\mathrm{e}(f_2,g_2).
\end{aligned}$$

This proves (i), and (ii) follows from (i) and Lemma 6.3. The proof of the lemma is finished.

Lemma 7.5. *For any f and g in $\mathscr{P}_2$, we have*

$$\mathrm{e}(f^{(n)},g^{(n)})\leqq\mathrm{e}(f,g),\qquad n\geqq 1. \tag{7.7}$$

Proof. Since (7.7) is evident for $n=1$, it is enough to prove that (7.7) holds for $n=m$ assuming that it holds for $n<m$. Making use of Lemma 6.3 first and then (ii) of Lemma 7.4, we have

$$\begin{aligned}
\mathrm{e}(f^{(m)},g^{(m)})&=\mathrm{e}\left(\frac{1}{m-1}\sum_{k=1}^{m-1}f^{(k)}\circ f^{(m-k)},\frac{1}{m-1}\sum_{k=1}^{m-1}g^{(k)}\circ g^{(m-k)}\right)\\
&\leqq\frac{1}{m-1}\sum_{k=1}^{m-1}\mathrm{e}(f^{(k)}\circ f^{(m-k)},g^{(k)}\circ g^{(m-k)})\\
&\leqq\frac{1}{m-1}\sum_{k=1}^{m-1}\{\gamma\,\mathrm{e}(f^{(k)},g^{(k)})+(1-\gamma)\,\mathrm{e}(f^{(m-k)},g^{(m-k)})\}\\
&\leqq\mathrm{e}(f,g),
\end{aligned}$$

as was to be proved.

Now the proof of the theorem in *Case I* is completed as follows. (7.3) is immediate from (i) of Lemma 7.4. To prove (7.4), we notice that

$$T_{t+s}f=e^{-qs}\sum_{n=1}^{\infty}(1-e^{-qs})^{n-1}(T_tf)^{(n)},$$

$$T_{t+s}g=e^{-qs}\sum_{n=1}^{\infty}(1-e^{-qs})^{n-1}(T_tg)^{(n)},$$

and then apply Lemma 6.3 and 7.5. The result is

$$\begin{aligned}&\mathrm{e}(T_{t+s}f, T_{t+s}g)\\ &\quad\leqq e^{-qs}\,\mathrm{e}(T_tf, T_tg)+e^{-qs}(1-e^{-qs})\,\mathrm{e}[(T_tf)^{(2)}, (T_tg)^{(2)}]\\ &\qquad+(1-2e^{-qs}+e^{-2qs})\,\mathrm{e}(T_tf, T_tg),\end{aligned}$$

and hence

$$\begin{aligned}&\overline{\lim_{s\downarrow 0}}\,\{\mathrm{e}(T_{t+s}f, T_{t+s}g)-\mathrm{e}(T_tf, T_tg)\}/s\\ &\quad\leqq -q\,\{\mathrm{e}(T_tf, T_tg)-\mathrm{e}[(T_tf)^{(2)}, (T_tg)^{(2)}]\}\\ &\quad\leqq -q(2\pi/q)\int_0^\pi \{\mathrm{e}(T_tf, T_tg)-\mathrm{e}[[T_tf]_\theta, [T_tg]_\theta]\}\,Q(d\theta)\\ &\quad= -2\pi\int_0^\pi \bar{\mathrm{e}}_\theta(T_tf, T_tg)\,Q(d\theta).\end{aligned}$$

The inequality (7.4) now follows from the above, since $\mathrm{e}\,(T_tf, T_tg)$ is continuous in t. The proof in *Case I* is finished.

Case II. We deal with the case when $\int_0^\pi Q(d\theta)=\infty$. For each $\varepsilon\in(0,\pi)$, the result in *Case I* is applicable to the semigroup $\{T_t^{(\varepsilon)}\}$ which is associated with $Q_\varepsilon(d\theta)$, and hence

$$\mathrm{e}(T_t^{(\varepsilon)}f, T_t^{(\varepsilon)}g)\leqq \mathrm{e}(f,g)-2\pi\int_0^t ds\int_\varepsilon^\pi \bar{\mathrm{e}}_\theta(T_s^{(\varepsilon)}f, T_s^{(\varepsilon)}g)\,Q(d\theta). \tag{7.8}$$

On the other hand, making use of (i) of Lemma 7.4 and (ii) of Lemma 7.2, we have $\rho[[T_s^{(\varepsilon)}f]_\theta, [T_sf]_\theta]\leqq\rho(T_s^{(\varepsilon)}f, T_sf)\to 0$ as $\varepsilon\downarrow 0$ and hence

$$\rho[[T_s^{(\varepsilon)}f]_\theta, [T_s^{(\varepsilon)}g]_\theta]\to\rho[[T_sf]_\theta, [T_sg]_\theta], \qquad \varepsilon\downarrow 0.$$

Therefore we have $\bar{\mathrm{e}}_\theta(T_s^{(\varepsilon)}f, T_s^{(\varepsilon)}g)\to\bar{\mathrm{e}}_\theta(T_sf, T_sg)$ as $\varepsilon\downarrow 0$, the convergence being bounded. Now, letting $\varepsilon\downarrow 0$ in (7.8) we obtain (7.4). Thus the proof of the theorem is completed.

§ 8. Theorem of Ikenberry and Truesdell on Time Evolution of Moments

The result on the time evolution of the moments for solutions to Boltzmann's equation of Maxwellian molecules goes back to Ikenberry and Truesdell [4]. In [4], however, the existence of solutions of the Boltzmann equation is not discussed rigorously. Here we state and prove the theorem of Ikenberry and Truesdell in our setting, for completeness. We state also a corollary; this will be useful in the next section where a more precise result on the trend to equilibrium will be obtained in connection with our metric ρ.

The method of [4] is to use harmonic polynomials. For each $k\geqq 0$ we choose $2k+1$ linearly independent (homogeneous) harmonic polynomials $\{\xi_k^l(x)\}_{|l|\leqq k}$ of degree k in $\mathbf{R}^3$ and put

$$\xi_{\mathbf{n}}(x)=|x|^{2r}\,\xi_k^l(x) \qquad \text{for } \mathbf{n}=(r,k,l),$$
$$\xi_{(0,0,0)}(x)=1,$$

where $r=0,1,\ldots,k=0,1,\ldots$, and $|l|\leqq k$. The degree of $\xi_{\mathbf{n}}$ is $|\mathbf{n}|=2r+k$. Then it is well-known that any homogeneous polynomial of x with degree n can be expressed by a linear combination of $\xi_{\mathbf{n}}(x)$ with $|\mathbf{n}|=n$, and therefore when dealing with moments of a probability distribution f on $\mathbf{R}^3$ it is sufficient to consider only the (harmonic) moments $\mu_{\mathbf{n}}(f)\equiv\langle f,\xi_{\mathbf{n}}\rangle$.

Lemma 8.1. (i) *Let $h(x)$, $|x|=1$, be a spherical harmonic of degree k and y be a unit vector in $\mathbf{R}^3$. On the unit sphere $S^2=\{|x|=1\}$ we take a spherical coordinate system with polar axis y, and denote by γ and ψ the colatitude and the longitude, respectively, of a point $x\in S^2$. Then*

$$\int_0^{2\pi} h(x)\,d\psi=2\pi\,P_k(\cos\gamma)\,h(y), \tag{8.1}$$

where P_k denotes the Legendre polynomial of degree k.

(ii) *If $\xi(x)$ is a (homogeneous) harmonic polynomial of degree k, then*

$$\langle\Pi_{x,-x,\theta},\xi\rangle=P_k(\cos\theta)\,\xi(x).$$

Proof. (i) is known as the mean value theorem for spherical harmonics; for the proof it is enough to check (8.1) for each h in the list

$$\begin{aligned}&P_k(\cos\gamma)\\&P_k^{(m)}(\cos\gamma)\sin^m\gamma\cos m\psi\\&P_k^{(m)}(\cos\gamma)\sin^m\gamma\sin m\psi,\qquad m=1,\ldots,k,\end{aligned} \tag{8.2}$$

because (8.2) forms a basis for the vector space of spherical harmonics of degree k ([3]). (ii) follows from (i), since ξ can be expressed as $\xi(x)=|x|^k\,h(x/|x|)$, $x\in\mathbf{R}^3$, with some spherical harmonic h of degree k.

Theorem 8.1 (Ikenberry and Truesdell [4]). *Given a probability distribution f on $\mathbf{R}^3$ with $\int|x|^\nu f(dx)<\infty$ for some integer $\nu\geqq 1$, we put $\mu_{\mathbf{n}}(t)=\mu_{\mathbf{n}}(T_tf)$ for $\mathbf{n}$ such that $|\mathbf{n}|\leqq\nu$. Then, for any $\mathbf{n}$ with $|\mathbf{n}|\leqq\nu$ we have*

$$\frac{d}{dt}\mu_{\mathbf{n}}(t)=\textstyle\sum'\beta^{\mathbf{n}}_{\mathbf{n}_1,\mathbf{n}_2}\,\mu_{\mathbf{n}_1}(t)\,\mu_{\mathbf{n}_2}(t)-\beta_{\mathbf{n}}\,\mu_{\mathbf{n}}(t), \tag{8.3}$$

where $\sum'$ means the summation taken over all pairs $(\mathbf{n}_1,\mathbf{n}_2)$ satisfying $|\mathbf{n}_1|+|\mathbf{n}_2|=|\mathbf{n}|$ and $|\mathbf{n}_1|,|\mathbf{n}_2|\geqq 1$. $\beta_{\mathbf{n}}$ is given by

$$\beta_{\mathbf{n}}=2\pi\int_0^\pi\left\{1-\left(\cos\frac{\theta}{2}\right)^{|\mathbf{n}|}P_k\left(\cos\frac{\theta}{2}\right)-\left(\sin\frac{\theta}{2}\right)^{|\mathbf{n}|}P_k\left(\sin\frac{\theta}{2}\right)\right\}Q(d\theta)>0,$$

and $\beta^{\mathbf{n}}_{\mathbf{n}_1,\mathbf{n}_2}$ are also some constants.

Proof. The proof will be given in four steps.

Step 1. We prove that

$$\langle \Pi_{x,0,\theta}, \xi_{\mathbf{n}} \rangle = \left(\cos\frac{\theta}{2}\right)^{|\mathbf{n}|} P_k\left(\cos\frac{\theta}{2}\right)\xi_{\mathbf{n}}(x), \qquad \mathbf{n}=(r,k,l).$$

In fact, from the relation $\Pi_{x,0,\theta} = \Pi_{x\cos(\theta/2),\,-x\cos(\theta/2),\,\theta/2}$ it follows that

$$\begin{aligned}\langle \Pi_{x,0,\theta}, \xi_{\mathbf{n}} \rangle &= \langle \Pi_{x\cos(\theta/2),\,-x\cos(\theta/2),\,\theta/2}, \xi_{\mathbf{n}} \rangle \\ &= |x|^{2r}\left(\cos\frac{\theta}{2}\right)^{2r} \langle \Pi_{x\cos(\theta/2),\,-x\cos(\theta/2),\,\theta/2}, \xi_k^l \rangle \\ &= \left(\cos\frac{\theta}{2}\right)^{|\mathbf{n}|} P_k\left(\cos\frac{\theta}{2}\right)\xi_{\mathbf{n}}(x).\end{aligned}$$

Step 2 is to prove that

$$\begin{aligned}K_{\mathbf{n}}(x,y) &\equiv 2\pi\int_0^\pi \{\langle \Pi_{x,y,\theta}, \xi_{\mathbf{n}} \rangle - \xi_{\mathbf{n}}(x)\}\, Q(d\theta) \\ &= (K\xi_{\mathbf{n}})(x,y)\end{aligned}$$

is a homogeneous polynomial in x and y of degree $|\mathbf{n}|$. Since we can write $\xi_{\mathbf{n}}(y+x) = \sum_i \eta_i(x)\zeta_i y)$, where η_i and ζ_i are homogeneous polynomials of degree m_i and n_i, respectively, with $m_i + n_i = |\mathbf{n}|$, we have

$$\begin{aligned}\langle \Pi_{x,y,\theta}, \xi_{\mathbf{n}} \rangle &= \langle \Pi_{x-y,0,\theta}, \xi_{\mathbf{n}}(y+\cdot) \rangle \\ &= \sum_i \zeta_i(y) \langle \Pi_{x-y,0,\theta}, \eta_i \rangle.\end{aligned}$$

On the other hand, η_i can be expressed as

$$\eta_i = \sum_{|\mathbf{m}|=m_i} c^i_{\mathbf{m}} \xi_{\mathbf{m}}, \qquad \mathbf{m}=(s,j,l'),$$

and hence from Step 1 we have

$$\langle \Pi_{x,y,\theta}, \xi_{\mathbf{n}} \rangle = \sum_i \zeta_i(y) \sum_{|\mathbf{m}|=m_i} c^i_{\mathbf{m}} \left(\cos\frac{\theta}{2}\right)^{m_i} P_j\left(\cos\frac{\theta}{2}\right)\xi_{\mathbf{m}}(x-y).$$

Therefore

$$\begin{aligned}&\langle \Pi_{x,y,\theta}, \xi_{\mathbf{n}} \rangle - \xi_{\mathbf{n}}(x) \\ &\quad = \sum_i \sum_{|\mathbf{m}|=m_i} \left\{\left(\cos\frac{\theta}{2}\right)^{m_i} P_j\left(\cos\frac{\theta}{2}\right) - 1\right\} c^i_{\mathbf{m}} \zeta_i(y)\,\xi_{\mathbf{m}}(x-y).\end{aligned}$$

This is a homogeneous polynomial in x and y of degree $|\mathbf{n}|$ with coefficients depending upon θ in such a way that they are $O(\theta)$ as $\theta \downarrow 0$. Thus $K_{\mathbf{n}}(x,y)$ is a homogeneous polynomial in x and y of degree $|\mathbf{n}|$.

Step 3 is to prove that

$$K_{\mathbf{n}}(x,0)=-\beta'_{\mathbf{n}}\xi_{\mathbf{n}}(x),\qquad K_{\mathbf{n}}(0,y)=-\beta''_{\mathbf{n}}\xi_{\mathbf{n}}(y) \tag{8.4}$$

where

$$\beta'_{\mathbf{n}}=2\pi\int_0^\pi\left\{1-\left(\cos\frac{\theta}{2}\right)^{|\mathbf{n}|}P_k\left(\cos\frac{\theta}{2}\right)\right\}Q(d\theta),$$

$$\beta''_{\mathbf{n}}=-2\pi\int_0^\pi\left(\sin\frac{\theta}{2}\right)^{|\mathbf{n}|}P_k\left(\sin\frac{\theta}{2}\right)Q(d\theta),\qquad \mathbf{n}=(r,k,l).$$

In fact, the first expression of (8.4) follows immediately from Step 1. As for the second, noting $\Pi_{0,y,\theta}=\Pi_{y,0,\pi-\theta}$ and then using the result of Step 1 we have

$$\begin{aligned}\langle\Pi_{0,y,\theta},\xi_{\mathbf{n}}\rangle&=\langle\Pi_{y,0,\pi-\theta},\xi_{\mathbf{n}}\rangle\\&=\left(\cos\frac{\pi-\theta}{2}\right)^{|\mathbf{n}|}P_k\left(\cos\frac{\pi-\theta}{2}\right)\xi_{\mathbf{n}}(y)\\&=\left(\sin\frac{\theta}{2}\right)^{|\mathbf{n}|}P_k\left(\sin\frac{\theta}{2}\right)\xi_{\mathbf{n}}(y).\end{aligned}$$

This implies the second expression of (8.4).

Step 4. If we set $J_{\mathbf{n}}(x,y)=K_{\mathbf{n}}(x,y)-K_{\mathbf{n}}(x,0)-K_{\mathbf{n}}(0,y)$, then by Step 2 the polynomial $J_{\mathbf{n}}(x,y)$ consists only of those terms which have at least degree 1 in x as well as in y, and therefore it can be expressed as

$$J_{\mathbf{n}}(x,y)=\sum{}'\beta^{\mathbf{n}}_{\mathbf{n}_1,\mathbf{n}_2}\xi_{\mathbf{n}_1}(x)\xi_{\mathbf{n}_2}(y).$$

Now, keeping in mind the moment estimate (4.30), we obtain

$$\begin{aligned}\frac{d}{dt}\mu_{\mathbf{n}}(t)&=\langle(T_tf)\otimes(T_tf),K_{\mathbf{n}}(x,y)\rangle\\&=\langle(T_tf)\otimes(T_tf),J_{\mathbf{n}}(x,y)\rangle\\&\quad+\langle(T_tf)\otimes(T_tf),K_{\mathbf{n}}(x,0)+K_{\mathbf{n}}(0,y)\rangle\\&=\sum{}'\beta^{\mathbf{n}}_{\mathbf{n}_1,\mathbf{n}_2}\mu_{\mathbf{n}_1}(t)\mu_{\mathbf{n}_2}(t)-\beta_{\mathbf{n}}\mu_{\mathbf{n}}(t),\end{aligned}$$

where $\beta_{\mathbf{n}}=\beta'_{\mathbf{n}}+\beta''_{\mathbf{n}}$. This completes the proof of the theorem.

It should be noticed that, in the right hand side of (8.3), there appear only the moments of degree less than $|\mathbf{n}|$ except for $\mu_{\mathbf{n}}(t)$ and that the coefficient $\beta_{\mathbf{n}}$ of $\mu_{\mathbf{n}}(t)$ is positive (we exclude the trivial case $Q\equiv 0$). As a consequence we have the following corollary which is also found in [4].

Corollary. *Let f be a probability distribution on $\mathbf{R}^3$ with finite absolute moments of all degrees, and assume that*

$$\int|x-m|^2f(dx)=3v>0,\qquad m=\int xf(dx). \tag{8.5}$$

Let g be the Gaussian distribution

$$(2\pi v)^{-3/2}\exp(-|x-m|^2/2v)\,dx \tag{8.6}$$

and put $\mu_{\mathbf{n}} = \mu_{\mathbf{n}}(\mathfrak{g})$. *Then, for each* $\mathbf{n}$

$$\mu_{\mathbf{n}}(t) \text{ converges to } \mu_{\mathbf{n}} \text{ exponentially fast as } t \to \infty. \tag{8.7}$$

In particular, $T_t f$ *converges to* $\mathfrak{g}$ *as* $t \to \infty$.

Proof. Clearly (8.7) holds for $|\mathbf{n}| = 0$ and 1 (the case of $|\mathbf{n}| = 1$ is nothing but (4.31)). So we assume that (8.7) holds for $0 \leq |\mathbf{n}| < k$ and prove it for $|\mathbf{n}| = k$. First we notice that $\mathfrak{g}$ is invariant under T_t (see **2** of Appendix). This implies

$$\sum{}' \beta^{\mathbf{n}}_{\mathbf{n}_1, \mathbf{n}_2} \mu_{\mathbf{n}_1} \mu_{\mathbf{n}_2} - \beta_{\mathbf{n}} \mu_{\mathbf{n}} = 0. \tag{8.8}$$

For any $\mathbf{n}$ with $|\mathbf{n}| = k$ we put $\bar{\mu}_{\mathbf{n}}(t) = \sum' \beta^{\mathbf{n}}_{\mathbf{n}_1, \mathbf{n}_2} \mu_{\mathbf{n}_1}(t) \mu_{\mathbf{n}_2}(t)$. Then the induction hypothesis implies that $\bar{\mu}_{\mathbf{n}}(t)$ converges, exponentially fast as $t \to \infty$, to $\sum' \beta^{\mathbf{n}}_{\mathbf{n}_1, \mathbf{n}_2} \mu_{\mathbf{n}_1} \mu_{\mathbf{n}_2}$ which is equal to $\beta_{\mathbf{n}} \mu_{\mathbf{n}}$ by (8.8). This fact combined with (8.3) implies that

$$\mu_{\mathbf{n}}(t) = e^{-\beta_{\mathbf{n}} t} \mu_{\mathbf{n}}(0) + \int_0^t e^{-\beta_{\mathbf{n}}(t-s)} \bar{\mu}_{\mathbf{n}}(s)\, ds \to \mu_{\mathbf{n}}, \quad \text{exponentially fast as } t \to \infty.$$

So the proof is finished.

§ 9. Proof of the Trend to Equilibrium

In this section we make use of the results of § 7 to prove the trend to equilibrium without assuming the existence of higher absolute moments. Fundamentally, our theorem is an extension of the result [17] in Kac's one-dimensional model to the case of Boltzmann's equation of Maxwellian molecules.

Theorem 9.1. *Let* $f \in \mathscr{P}_2$ *and assume that* (8.5) *is satisfied. Let* $\mathfrak{g}$ *be the Gaussian distribution* (8.6) *and put* $\mathfrak{e}(f) = \mathfrak{e}(f, \mathfrak{g})$. *Then,* $\mathfrak{e}(T_t f)$ *decreases to* 0 *as* $t \uparrow \infty$. *In particular,* $T_t f$ *converges to* $\mathfrak{g}$ *as* $t \uparrow \infty$.

The proof is based on the following lemma.

Lemma 9.1. *Let* f *and* $\mathfrak{g}$ *be the same as in the theorem and put* $\bar{\mathfrak{e}}_\theta(f) = \bar{\mathfrak{e}}_\theta(f, \mathfrak{g})$. *Then,* $\bar{\mathfrak{e}}_\theta(f) > 0$ *for* $0 < \theta < \pi$ *if* $f \neq \mathfrak{g}$.

Proof of the Lemma. Since $(\mathfrak{g} \circ \mathfrak{g})_\theta = \mathfrak{g}$ by **2** of Appendix, what we have to prove is $\mathfrak{e}[(f \circ f)_\theta, (\mathfrak{g} \circ \mathfrak{g})_\theta] < \mathfrak{e}(f, \mathfrak{g})$ for $0 < \theta < \pi$ provided $f \neq \mathfrak{g}$. By (i) of Lemma 7.4 we have

$$\mathfrak{e}[(f \circ f)_\theta, (\mathfrak{g} \circ \mathfrak{g})_\theta] \leq \mathfrak{e}(f, \mathfrak{g}). \tag{9.1}$$

So, assuming the equality holds in the above, we prove that $f = \mathfrak{g}$. We now recall the proof of (i) of Lemma 7.4. Then, in order to have the equality in (9.1), the inequality (7.6) must be the equality

$$(X_1 - X_2, Y_1 - Y_2) = |X_1 - X_2| |Y_1 - Y_2|, \quad \text{a.s.,}$$

which is equivalent to

$$\frac{X_1 - X_2}{|X_1 - X_2|} = \frac{Y_1 - Y_2}{|Y_1 - Y_2|}, \quad \text{a.s.,} \tag{9.2}$$

On the other hand, both Y_1 and Y_2 are $\mathfrak{g}$-distributed in the present case, by Theorem 1 of [12] (or by (6.2)) there exists a unique Borel mapping ψ from $\mathbf{R}^3$ into itself such that $X_i=\psi(Y_i)$, $i=1,2$, almost surely (the uniqueness of ψ was remarked in the proof of Lemma 6.2). Therefore, (9.2) yields

$$\frac{\psi(y_1)-\psi(y_2)}{|\psi(y_1)-\psi(y_2)|}=\frac{y_1-y_2}{|y_1-y_2|}$$

for almost all $y_1, y_2 \in \mathbf{R}^3$ with respect to the Lebesgue measure. Thus for some $y_0 \in \mathbf{R}^3$

$$\psi(y)=\psi(y_0)+\frac{|\psi(y)-\psi(y_0)|}{|y-y_0|}(y-y_0) \tag{9.3}$$

must hold for almost all y. Therefore,

$$\frac{|\psi(y_1)-\psi(y_0)|}{|y_1-y_0|}(y_1-y_0)-\frac{|\psi(y_2)-\psi(y_0)|}{|y_2-y_0|}(y_2-y_0)$$
$$=\psi(y_1)-\psi(y_2)=\frac{|\psi(y_1)-\psi(y_2)|}{|y_1-y_2|}(y_1-y_2), \quad \text{a.e.,}$$

and hence

$$\frac{|\psi(y_1)-\psi(y_0)|}{|y_1-y_0|}=\frac{|\psi(y_2)-\psi(y_0)|}{|y_2-y_0|}, \quad \text{a.e.}$$

This combined with (9.3) implies that $\psi(y)=\psi(y_0)+const\,(y-y_0)$, a.e., and hence $\psi(y)=y$ because $E\{X_1\}=E\{Y_1\}$ and $E\{|X_1|^2\}=E\{|Y_1|^2\}$. Thus we obtain $f=\mathfrak{g}$, as was to be proved.

Proof of the Theorem. Since $\mathfrak{e}(T_tf, T_t\mathfrak{g})=\mathfrak{e}(T_tf,\mathfrak{g})=\mathfrak{e}(T_tf)$, the decreasing property of $\mathfrak{e}(T_tf)$ in t follows from Theorem 7.1. To prove that $\mathfrak{e}(T_tf)$ tends to 0 as $t\uparrow\infty$, we first assume that $\int|x|^4 f(dx)<\infty$. Then by the corollary to Theorem 8.1 $\int|x|^4(T_tf)(dx)$ is bounded in t, say, by M. We denote by $\tilde{\mathscr{P}}$ the family of probability distributions $\tilde{f}$ on $\mathbf{R}^3$ satisfying

$$\int x\tilde{f}(dx)=m, \qquad \int|x-m|^2\tilde{f}(dx)=3v, \qquad \int|x|^4\tilde{f}(dx)\leqq M,$$

and put $\tilde{\mathscr{P}}_\varepsilon=\{\tilde{f}\in\tilde{\mathscr{P}}:\mathfrak{e}(\tilde{f})\geqq\varepsilon\}$ for $\varepsilon>0$. Then $\tilde{\mathscr{P}}_\varepsilon$ is compact with respect to the metric ρ. Moreover, using the triangle inequality for ρ and (i) of Lemma 7.4 we can see that $|\bar{\mathfrak{e}}_\theta(\tilde{f})-\bar{\mathfrak{e}}_\theta(\tilde{g})|\leqq 2\mathfrak{e}(\tilde{f},\tilde{g})$ for $\tilde{f},\tilde{g}\in\tilde{\mathscr{P}}_\varepsilon$ and hence $\bar{\mathfrak{e}}_\theta$ is ρ-continuous on $\tilde{\mathscr{P}}_\varepsilon$ for each $\theta\in(0,\pi)$. Since $\bar{\mathfrak{e}}_\theta$ is strictly positive on $\tilde{\mathscr{P}}_\varepsilon$ by Lemma 9.1, we have

$$\inf_{\tilde{f}\in\tilde{\mathscr{P}}_\varepsilon}\bar{\mathfrak{e}}_\theta(\tilde{f})>0, \qquad \theta\in(0,\pi),$$

and hence

$$\Phi(\varepsilon)\equiv\inf_{\tilde{f}\in\tilde{\mathscr{P}}_\varepsilon}2\pi\int_0^\pi\bar{\mathfrak{e}}_\theta(\tilde{f})\,Q(d\theta)>0. \tag{9.4}$$

On the other hand, from (7.4) we have

$$e(T_t f) \leqq e(f) - 2\pi \int_0^t ds \int_0^\pi \bar{e}_\theta(T_s f)\, Q(d\theta).$$

Because $T_t f \in \tilde{\mathscr{P}}_\varepsilon$ with $\varepsilon = e(T_t f)$ if $e(T_t f) > 0$, the above inequality combined with (9.4) implies

$$e(T_t f) \leqq e(f) - \int_0^t \Phi[e(T_s f)]\, ds \tag{9.5}$$

for t such that $e(T_t f) > 0$. But, this inequality clearly implies that $e(T_t f) \to 0$ as $t \to \infty$.

Finally we remove the restriction $\int |x|^4 f(dx) < \infty$. For each $\varepsilon > 0$ and $f \in \mathscr{P}_2$ satisfying (8.5) we can choose f_ε in such a way that

$$\int x f_\varepsilon(dx) = m, \qquad \int |x-m|^2 f_\varepsilon(dx) = 3v, \qquad \int |x|^4 f_\varepsilon(dx) < \infty,$$

and $\rho(f, f_\varepsilon) < \varepsilon$ hold. Then, using Theorem 8.1 we have

$$\begin{aligned} e(T_t f) &\leqq \{\rho(T_t f, T_t f_\varepsilon) + \rho(T_t f_\varepsilon, g)\}^2 \\ &\leqq \{\varepsilon + \sqrt{e(T_t f_\varepsilon)}\}^2 \end{aligned}$$

and hence $\overline{\lim}_{t \to \infty} e(T_t f) \leqq \varepsilon^2$. The proof of the theorem is completed.

Remark. Making use of the corollary to Theorem 8.1 in full, we obtain a much simpler proof of Theorem 9.1. If f has finite absolute moments of all degrees, then $\mu_{\mathbf{n}}(T_t f) \to \mu_{\mathbf{n}}$ as $t \to \infty$ for every $\mathbf{n}$ and hence $e(T_t f) \to 0$ as $t \to \infty$. The general case when f belongs to $\mathscr{P}_2$ and satisfies (8.5) can be treated by choosing f_ε with finite absolute moments of all degrees in such a way that $\int x f_\varepsilon(dx) = m$, $\int |x-m|^2 f_\varepsilon(dx) = 3v$ and $\rho(f, f_\varepsilon) < \varepsilon$ hold. However, our first proof based upon the inequality (9.5) seems to be interesting.

Appendix

1. In the introduction we regarded the equation (0.3) as a weak version of (0.2). This is justified by the formula

$$\begin{aligned} &\int_{(0,\pi)\times(0,2\pi)\times\mathbf{R}^6} \xi(x)(u'u'_1 - uu_1)\, Q(d\theta)\, d\varphi\, dx\, dx_1 \\ &\quad = \int_{\mathbf{R}^6} (K\xi)(x, x_1)\, u(x)\, u(x_1)\, dx\, dx_1, \qquad \xi \in C_0^\infty(\mathbf{R}^3), \end{aligned}$$

which holds, at least if $u(x)$ is smooth enough, according to the following lemma.

Lemma. *Let $\xi(x, x_1, y, y_1)$ be a continuous function on R^{12} with compact support. Then, for each* $0 \in (0, \pi)$

$$\int_{(0,2\pi)\times\mathbf{R}^6} \xi(x, x_1, x', x'_1)\, d\varphi\, dx\, dx_1 = \int_{(0,2\pi)\times\mathbf{R}^6} \xi(x', x'_1, x, x_1)\, d\varphi\, dx\, dx_1. \tag{1}$$

Proof. We denote by $\Phi(\theta)$ and $\Psi(\theta)$ the left and right hand sides of (1), respectively. Since Φ and Ψ are continuous in θ, for the proof of (1) it is enough to show

$$\int_0^\pi \Phi(\theta) Q_0(\theta) \sin\theta\, d\theta = \int_0^\pi \Psi(\theta) Q_0(\theta) \sin\theta\, d\theta \tag{2}$$

for any continuous function $Q_0(\theta)$ with compact support in $(0, \pi)$. For $x \neq x_1$ we put $l = (x'-x)/|x'-x|$ and define $\gamma \in (0, \pi/2)$ by $\cos\gamma = (x_1 - x, l)/|x_1 - x|$. Since $\theta = \pi - 2\gamma$, $Q_0(\theta)\cos\gamma$ becomes a function of $|(x_1 - x, l)|$ and $|x_1 - x|$. Thus we can write $Q_0(\theta)\cos\gamma = F(x, x_1, l)$ with some function F on $\mathbf{R}^3 \times \mathbf{R}^3 \times S^2$ satisfying

$$F(x, x_1, l) = F(x', x_1', l) = F(x, x_1, -l). \tag{3}$$

We then have

$$\begin{aligned}
&\int_0^\pi \Phi(\theta) Q_0(\theta) \sin\theta\, d\theta \\
&\quad = 4 \int_{(0,\pi/2)\times(0,2\pi)\times\mathbf{R}^6} \xi(x, x_1, x', x_1') Q_0(\theta) \cos\gamma \sin\gamma\, d\gamma\, d\varphi\, dx\, dx_1 \\
&\quad = 4 \int_{\mathbf{R}^6} dx\, dx_1 \int_{\{l\in S^2:\, (x_1 - x,\, l) > 0\}} \xi(x, x_1, x', x_1') F(x, x_1, l)\, dl \\
&\quad = 2 \int_{\mathbf{R}^6\times S^2} \xi(x, x_1, x', x_1') F(x, x_1, l)\, dx\, dx_1\, dl;
\end{aligned} \tag{4}$$

in the last line of the above x' and x_1' are defined by

$$\begin{aligned}
x' &= x + (x_1 - x, l)\, l \\
x_1' &= x_1 - (x_1 - x, l)\, l.
\end{aligned} \tag{5}$$

Since $dx\, dx_1 = dx'\, dx_1'$ for each fixed $l \in S^2$, the last integral in (4) is equal to

$$\int_{\mathbf{R}^6\times S^2} \xi(x, x_1, x', x_1') F(x, x_1, l)\, dx'\, dx_1'\, dl, \tag{6}$$

where x and x_1 are defined by the same rule as (5):

$$\begin{aligned}
x &= x' + (x_1' - x', l)\, l \\
x_1 &= x_1' - (x_1' - x', l)\, l.
\end{aligned}$$

Now, from (3) it is clear that (6) is equal to

$$\int_{\mathbf{R}^6\times S^2} \xi(x', x_1', x, x_1) F(x, x_1, l)\, dx\, dx_1\, dl.$$

But this is equal to the right hand side of (2) by the same reason as (4) holds. The proof is finished.

2. Let g be the Gaussian distribution (8.6). Then g is invariant under T_t. For the proof, we first notice that the density function g satisfies $g' g_1' = g g_1$ and hence by the above lemma

$$\langle \mathrm{g} \otimes \mathrm{g}, K\xi \rangle = 0, \qquad \xi \in C_0^\infty(\mathbf{R}^3).$$

This implies $T_t g = g$ at least if $\int_0^\pi Q(d\theta) < \infty$, because the uniqueness of the solution for (0.3) clearly holds in that case. Therefore, in general, we have $T_t\, g = \lim_{\varepsilon \downarrow 0} T_t^{(\varepsilon)} g = g$, proving the invariance of g under T_t. Moreover, $(g \circ g)_\theta = g$ also follows from $g' g'_1 = g g_1$ and the above lemma.

References

1. Arkeryd, L.: On the Bolztmann equation, Part 1: Existence, Part 11: The full initial value problem. Arch Rational Mech. Anal. **45**, 1–16, 17–34 (1972)
2. Carleman, T.: Problèmes Mathèmatiques dans la Thèorie Cinètique des Gaz. Uppsala 1957.
3. Erdélyi, A. (Editor): Higher Transcendental Functions, Vol. 11, New York-Toronto-London: McGraw-Hill 1953
4. Ikenberry, E., Truesdell, C.: On the pressure and the flux of energy in a gas according to Maxwell's kinetic theory 1. J. Rational Mech. Anal. **5**, 1–54 (1956)
5. Itô, K.: On stochastic differential eqations. Mem. Amer. Math. Soc. No. 4, 1951.
6. Kac, M.: Probability and Related Topics in the Physical Sciences, New York-London: Interscience 1959.
7. Kolmogorov, A.N.: On Skorohod convergence. Theor. Probability Appl. **1**, 215–222 (1956)
8. Kondô, R., Negoro, A.: Certain functional of probability mesures on Hilbert spaces. Hiroshima Math. J. **6**, 421–428 (1976)
9. McKean, H.P.: A class of Markov processes associated with nonlinear parabolic equations. Proc. Nat. Acad. Sci. **56**, 1907–1911 (1966)
10. McKean, H.P.: An exponential formula for solving Boltzmann's equation for a Maxwellian gas. J. Combinatorial Theory **2**, 358–382 (1967)
11. McKean, H.P.: Propagation of chaos for a class of nonlinear parabolic equation. Lecture series in Differential Equtions, session 7: Catholic Univ. 1967
12. Murata, H., Tanaka, H.: An inequality for certain functional of multidimensional probability distributions. Hiroshima Math. J. **4**, 75–81 (1974)
13. Murata, H.: Propagation of chaos for Boltzmann-like equation of non-cutoff type in the plane. Hiroshima Math. J. **7**, 479–515 (1977)
14. Parthasarathy, K.R.: Probability Measures on Metric Spaces, New York-London: Academic Press 1967
15. Povzner, A.Ya.: The Boltzmann equation the kinetic theory of gases. Mat. Sb. **58**, 65–86 (1962). (English Transl.) Amer. Math. Soc. Transl. **(2)47**, 193–216 (1965)
16. Skorohod, A.V.: Limit theorems for stochastic processes, Theor. Probability Appl. **1**, 261–290 (1956)
17. Tanaka, H.: An inequality for a functional of probability distributions and its application to Kac's one-dimensional model of a Maxwellian gas. Z. Wahrscheinlichkeitstheorie verw. Gebiete **27**, 47–52 (1973)
18. Tanaka, H.: On Markov process corresponding to Boltzmann's equation of Maxwellian gas. In: Proc. 2nd Japan-USSR Sympos. on Probab. Theory, Lecture Notes in Math., **330**, 478–489. Berlin-Heidelberg-New York: Springer-Verlag 1973
19. Tanaka, H.: On the uniqueness of Markov process associated with the Boltzmann equation of Maxwellian molecules. In: Proc. Intern. Sympos. on Stochastic Differential Equations, Kyoto, 1976
20. Uhlenbeck, G.E., Ford, G.W.: Lectures in Statistical Mechanics. Amer. Math. Soc., Providence, Rhode Island, 1963
21. Wild, E.: On Boltzmann's equation in the kinetic theory of gases. Proc. Cambridge Philos. Soc. **47**, 602–609 (1951)

Received December 19, 1977

HIROSHIMA MATH. J.
9 (1979), 163–177

Stochastic Differential Equations with Reflecting Boundary Condition in Convex Regions

Hiroshi TANAKA

(Received September 16, 1978)

§ 1. Introduction

A. V. Skorohod [4] considered a stochastic differential equation for a reflecting diffusion process on $\bar{D}=[0, \infty)$ (see also McKean [2] [3]). This is the simplest case among stochastic differential equations subject to boundary conditions and can be solved easily. The purpose of this paper is to show that the multi-dimensional version of Skorohod's equation is still easy to solve if we assume that the domain D is convex.

Skorohod's equation describing a reflecting Brownian path ξ on $\bar{D}=[0, \infty)$ is

$$\xi = w + \varphi, \tag{1.1}$$

where w is a standard Brownian path and ξ is to be found as a $\bar{D}$-valued continuous function under the condition that $\varphi(t)$ increases only when $\xi(t)=0$. The equation (1.1) has a unique solution not only for almost all Brownian paths but also for *all* continuous functions w with $w(0)\in\bar{D}$, and the solution is given by

$$\xi(t) = \begin{cases} w(t) & \text{for} \quad 0 \leq t \leq T, \\ w(t) - \inf\{w(s)\colon T \leq s \leq t\} & \text{for} \quad t > T, \end{cases} \tag{1.2}$$

where $T=\inf\{t>0\colon w(t)<0\}$. Our first problem is to consider a multi-dimensional version of the equation (1.1) assuming that D is a *convex* domain. Although we can not obtain an explicit formula for the solution like (1.2), we are able to construct the unique solution ξ for any $\mathbf{R}^d$-valued continuous function w with $w(0)\in\bar{D}$ and to prove that ξ depends continuously on w, *if* D is a convex domain in $\mathbf{R}^d$ satisfying certain condition (Theorem 2.1). This additional condition is automatically satisfied if D is bounded or $d=2$. This result will lead to a simple solution to our second problem which is concerned with a stochastic differential equation with (normal) reflection having variable coefficients similar to Skorohod's. The following may be stressed.

(i) The boundary does not need to be smooth as far as the domain is assumed to be convex.

(ii) The diffusion coefficients may degenerate (however, in this case the path of

the solution might not behave like an ordinary reflection).
Our results are roughly stated as follows. The convexity assumption for the domain makes the situation quite similar to that in the whole space; in fact, the existence of solutions will be obtained assuming only the bounded continuity of the coefficients (Theorem 4.2), and the pathwise uniqueness of solutions will be proved under the same regularity assumption on the coefficients as found in the work of Watanabe and Yamada [9] for the case of whole space (Theorem 4.3).

However, it is noted that our methods and results are restricted to the case of reflecting boundary condition; the convexity assumption will not simplify the situation in the case of general boundary conditions such as discussed by Ikeda [1], Watanabe [8], Stroock and Varadhan [6] and Tsuchiya [7].

§2. A deterministic problem

An $\mathbf{R}^d$-valued function $\varphi(t)=(\varphi^1(t),\dots, \varphi^d(t))$ defined on $\mathbf{R}_+=[0, \infty)$ is said to be of bounded variation for simplicity if the component functions are of bounded variation on each finite t-interval. Given such a function $\varphi(t)$ which is right continuous and $\varphi(0)=0$, we put

$$|\varphi|(t) = \text{the total variation of } \varphi \text{ on } [0, t]$$
$$= \sup \sum_k |\varphi(t_k) - \varphi(t_{k-1})|,$$

where the supremum is taken over all partitions: $0=t_0<t_1<\cdots<t_n=t$. $\varphi(t)$ can be expressed as

$$\varphi(t) = \int_0^t \boldsymbol{n}(s)d|\varphi|(s) = \int_{[0,t]} \boldsymbol{n}(s)\,|\varphi|(ds) \tag{2.1}$$

with a unit vector valued function $\boldsymbol{n}(t)$; $\boldsymbol{n}(t)$ is uniquely determined almost everywhere with respect to the measure $d|\varphi|$.

Let D be a convex domain in $\mathbf{R}^d$ and $\bar{D}$ its closure; D will be fixed throughout. For $x\in\partial D$ we denote by $\mathscr{H}_x(D)$ the set of all supporting hyperplanes of D at x. By (an inward) normal vector at $x\in\partial D$ we mean *any* inward unit vector perpendicular to some $H\in\mathscr{H}_x(D)$, and denote by $\mathscr{N}_x(D)$ the set of all inward normal vectors at $x\in\partial D$. Of course, it can happen that $\sharp\mathscr{N}_x(D)=\infty$ unless ∂D is smooth near x. We shall also consider the following spaces of functions.

$\mathbf{C}(\mathbf{R}_+, \mathbf{R}^d)$ (resp. $\mathbf{C}(\mathbf{R}_+, \bar{D})$) = the space of $\mathbf{R}^d$-valued (resp. $\bar{D}$-valued) continuous functions on $\mathbf{R}_+$.

$\mathbf{D}(\mathbf{R}_+, \mathbf{R}^d)$ (resp. $\mathbf{D}(\mathbf{R}_+, \bar{D})$) = the space of $\mathbf{R}^d$-valued (resp. $\bar{D}$-valued) right continuous functions on $\mathbf{R}_+$ with left limits.

On $\mathbf{C}(\mathbf{R}_+, \mathbf{R}^d)$ and $\mathbf{C}(\mathbf{R}_+, \bar{D})$ we consider the compact uniform topology. Given a function ξ in $\mathbf{D}(\mathbf{R}_+, \bar{D})$, a function φ is said to be *associated* with ξ if the following three conditions are satisfied.

(2.2) φ is a function in $\mathbf{D}(\mathbf{R}_+, \mathbf{R}^d)$ with bounded variation and $\varphi(0)=0$.

(2.3) The set $\{t \in \mathbf{R}_+ : \xi(t) \in D\}$ has $d|\varphi|$-measure zero.

(2.4) The function $\boldsymbol{n}(t)$ appearing in the expression (2.1) is a normal vector at $\xi(t)$ for almost all t with respect to the measure $d|\varphi|$.

REMARK 2.1. The condition (2.4) can be replaced by the following one.

(2.4′) For any $\eta \in \mathbf{C}(\mathbf{R}_+, \bar{D})$, $(\eta(t) - \xi(t), \varphi(dt)) \geq 0$.

EXAMPLE. Let ∂D be *smooth*, $\boldsymbol{n}(x)$ the inward normal vector at $x \in \partial D$ and $\xi \in \mathbf{D}(\mathbf{R}_+, \bar{D})$. Then, for any right continuous non-decreasing function $\rho(t)$ on $\mathbf{R}_+$ with $\rho(0)=0$

$$\varphi(t) = \int_0^t \mathbf{1}_{\partial D}(\xi(s))\boldsymbol{n}(\xi(s))d\rho(s)$$

is clearly an associated function of ξ.

Our first problem stated in the introduction can now be formulated as follows.

PROBLEM. Given $w \in \mathbf{D}(\mathbf{R}_+, \mathbf{R}^d)$ with $w(0) \in \bar{D}$, find a solution ξ of

$$\xi = w + \varphi. \tag{2.5}$$

When we speak of the equation (2.5), it is always understood that $\xi \in \mathbf{D}(\mathbf{R}_+, \bar{D})$ and φ is associated with ξ.

As stated in the introduction, in the simplest case $\bar{D} = [0, \infty)$ the solution of (2.5) is given by (1.2). However, in the general multi-dimensional case the existence of a solution of (2.5) is not trivial. An example in which a solution of (2.5) can easily be found is the case when w is a step function as will be seen in the lemma below. For a given point $x \in \mathbf{R}^d - \bar{D}$ we denote by $[x]_\partial$ the (unique) point on ∂D which gives the minimum distance between x and $\bar{D}$.

LEMMA 2.1. *If w is a step function with $w(0) \in \bar{D}$, then a solution of* (2.5) *exists.*

PROOF. Put $T_1 = \inf\{t>0 : w(t) \notin \bar{D}\}$ and define $\xi(t)$, $0 \leq t \leq T_1$, by $\xi(t) = w(t)$ for $t < T_1$ and $\xi(T_1) = [w(T_1)]_\partial$. Then, $T_1 > 0$ and $\xi(t)$ solves (2.5) for $0 \leq t \leq T_1$. Next, suppose a solution $\xi(t)$ of (2.5) is obtained for $0 \leq t \leq T_{n-1}$ and put

$$T_n = \inf\{t > T_{n-1}: w(t) + \varphi(T_{n-1}) \notin \bar{D}\},$$

$$\xi(t) = \begin{cases} w(t) + \varphi(T_{n-1}) & \text{for } T_{n-1} < t < T_n, \\ [w(T_n) + \varphi(T_{n-1})]_\partial & \text{for } t = T_n. \end{cases}$$

Then, $\xi(t)$ solves (2.5) for $0 \le t \le T_n$. Repeating this argument, we can obtain a solution of (2.5) for $0 \le t < \infty$ because $T_n \uparrow \infty$ as $n \uparrow \infty$.

LEMMA 2.2. (i) *Let* $w, \tilde{w} \in \mathbf{D}(\mathbf{R}_+, \mathbf{R}^d)$ *with* $w(0), \tilde{w}(0) \in \bar{D}$, *and* $\xi, \tilde{\xi}$ *be any solutions of*

$$\xi = w + \varphi, \quad \tilde{\xi} = \tilde{w} + \tilde{\varphi},$$

respectively. Then we have

$$|\xi(t) - \tilde{\xi}(t)|^2 \le |w(t) - \tilde{w}(t)|^2 + 2\int_0^t (w(t) - \tilde{w}(t) - w(s) + \tilde{w}(s), \varphi(ds) - \tilde{\varphi}(ds)).$$

(ii) *If* ξ *is a solution of* (2.5), *then*

$$|\xi(t) - \xi(s)|^2 \le |w(t) - w(s)|^2 + 2\int_{(s,t]} (w(t) - w(\tau), \varphi(d\tau)), \qquad 0 \le s \le t.$$

PROOF. (i) We have

$$\begin{aligned} |\varphi(t) - \tilde{\varphi}(t)|^2 &= \int_0^t\int_0^t (\varphi(dt_1) - \tilde{\varphi}(dt_1), \varphi(dt_2) - \tilde{\varphi}(dt_2)) \\ &= 2\iint_{0\le t_1 \le t_2 \le t} (\varphi(dt_1) - \tilde{\varphi}(dt_1), \varphi(dt_2) - \tilde{\varphi}(dt_2)) \\ &\quad - \sum_{0 \le s \le t} |\varphi(s) - \tilde{\varphi}(s) - \varphi(s-) + \tilde{\varphi}(s-)|^2 \\ &\le 2\int_0^t (\varphi(s) - \tilde{\varphi}(s), \varphi(ds) - \tilde{\varphi}(ds)), \end{aligned}$$

$$(w(t) - \tilde{w}(t), \varphi(t) - \tilde{\varphi}(t)) = \int_0^t (w(t) - \tilde{w}(t), \varphi(ds) - \tilde{\varphi}(ds))$$

$$= \int_0^t (w(t) - \tilde{w}(t) - w(s) + \tilde{w}(s), \varphi(ds) - \tilde{\varphi}(ds)) + \int_0^t (w(s) - \tilde{w}(s), \varphi(ds) - \tilde{\varphi}(ds)).$$

Therefore

$$|\xi(t) - \tilde{\xi}(t)|^2 = |w(t) - \tilde{w}(t)|^2 + 2(w(t) - \tilde{w}(t), \varphi(t) - \tilde{\varphi}(t)) + |\varphi(t) - \tilde{\varphi}(t)|^2$$

$$\leq |w(t) - \tilde{w}(t)|^2 + 2\int_0^t (\xi(s) - \tilde{\xi}(s), \varphi(ds) - \tilde{\varphi}(ds)) + 2\int_0^t (w(t) - \tilde{w}(t) - w(s) + \tilde{w}(s), \varphi(ds) - \tilde{\varphi}(ds)).$$

But the second term is non-positive by (2.4′) of Remark 2.1.

(ii) By a method similar to the case (i), we have

$$|\xi(t) - \xi(s)|^2 = |w(t) - w(s)|^2 + |\varphi(t) - \varphi(s)|^2 + 2(w(t) - w(s), \varphi(t) - \varphi(s))$$

$$\leq |w(t) - w(s)|^2 + 2\int_{(s,t]} (\xi(\tau) - \xi(s), \varphi(d\tau)) + 2\int_{(s,t]} (w(t) - w(\tau), \varphi(d\tau)).$$

But the second term is non-positive by (2.4′); the proof is finished.

REMARK 2.2. By a method similar to the above, we can prove the following: If w and $\tilde{w}$ of (i) of Lemma 2.2 are replaced by $w+a$ and $\tilde{w}+\tilde{a}$, respectively, where a and $\tilde{a}$ are $\mathbf{R}^d$-valued right continuous functions of bounded variation with $a(0)=\tilde{a}(0)=0$, then

$$|\xi(t) - \tilde{\xi}(t)|^2 \leq |w(t) - \tilde{w}(t)|^2 + 2\int_0^t (\xi(s) - \tilde{\xi}(s), a(ds) - \tilde{a}(ds)) + 2\int_0^t (w(t) - \tilde{w}(t) - w(s) + \tilde{w}(s), a(ds) + \varphi(ds) - \tilde{a}(ds) - \tilde{\varphi}(ds)).$$

By a similar replacement of w in (ii) by $w+a$, the inequality in (ii) becomes

$$|\xi(t) - \xi(s)|^2 \leq |w(t) - w(s)|^2 + 2\int_{(s,t]} (\xi(\tau) - \xi(s), a(d\tau)) + 2\int_{(s,t]} (w(t) - w(\tau), a(d\tau) + \varphi(d\tau)).$$

LEMMA 2.3. (2.5) *has at most one solution.*

PROOF. Let ξ and $\tilde{\xi}$ be solutions of (2.5). Then, putting $w=\tilde{w}$ in (i) of Lemma 2.2 we obtain $|\xi(t)-\tilde{\xi}(t)|^2 \leq 0$.

LEMMA 2.4. *If w is continuous, then the solution of* (2.5) *is also continuous.*

PROOF. From (ii) of Lemma 2.2 we have

$$|\xi(t)-\xi(s)|^2 \le |w(t)-w(s)|^2 + 2\int_{(s,t]} |w(t)-w(\tau)|\,|\varphi|(d\tau),$$

from which the continuity of ξ follows.

LEMMA 2.5. *Let $\{w_n\}_{n\ge 1}$ be a sequence in $\mathbf{D}(\mathbf{R}_+, \mathbf{R}^d)$ such that for each n the equation $\xi_n = w_n + \varphi_n$ has a solution for $0 \le t \le T$, T being a positive constant. If w_n converges uniformly on $[0, T]$ to some $w \in \mathbf{C}(\mathbf{R}_+, \mathbf{R}^d)$ as $n\to\infty$ and if $\{|\varphi_n|(T)\}_{n\ge 1}$ is bounded, then ξ_n converges uniformly on $[0, T]$ as $n\to\infty$ to the solution ξ of $\xi = w + \varphi$ for $0\le t\le T$.*

PROOF. Let K be the bound of $\{|\varphi_n|(T)\}_{n\ge 1}$. Then applying Lemma 2.2, we have

$$\text{(2.6a)}\qquad |\xi_n(t)-\xi_m(t)|^2 \le |w_n(t)-w_m(t)|^2 + 8K \sup_{0\le s\le t} |w_n(s)-w_m(s)|,$$

$$\text{(2.6b)}\qquad |\xi_n(t)-\xi_n(s)|^2 \le |w_n(t)-w_n(s)|^2 + 2K \sup_{s\le t_1\le t_2\le t} |w_n(t_2)-w_n(t_1)|.$$

From the first inequality it follows that $\{\xi_n\}_{n\ge 1}$ is uniformly convergent on $[0, T]$ and hence the same for $\{\varphi_n\}_{n\ge 1}$. Letting $n\uparrow\infty$ in the second inequality, we obtain the inequality concerning the limit functions ξ and φ:

$$|\xi(t)-\xi(s)|^2 \le |w(t)-w(s)|^2 + 2K \sup_{s\le t_1\le t_2\le t} |w(t_2)-w(t_1)|.$$

This implies the continuity of ξ. We now prove that ξ is a solution of (2.5) for $0\le t\le T$. For this it is enough to prove that φ is associated with ξ. First, $|\varphi_n|(T)\le K$ implies $|\varphi|(T)\le K$ and hence φ is of bounded variation. The condition (2.3) is also verified easily. To verify (2.4), let $\eta\in\mathbf{C}(\mathbf{R}_+, \bar{D})$ and notice that for $0\le t_1 < t_2\le T$

$$\left|\int_{t_1}^{t_2} (\eta(t)-\xi_n(t), \varphi_n(dt)) - \int_{t_1}^{t_2} (\eta(t)-\xi(t), \varphi(dt))\right|$$

$$\le \left|\int_{t_1}^{t_2} (\xi(t)-\xi_n(t), \varphi_n(dt))\right| + \left|\int_{t_1}^{t_2} (\eta(t)-\xi(t), \varphi_n(dt)-\varphi(dt))\right|.$$

The first is dominated by $K \sup_{t_1\le t\le t_2} |\xi(t)-\xi_n(t)|$ and hence tends to 0 as $n\to\infty$; the second term also does as can be seen by approximating the integral by the Riemann sum. Therefore

$$\int_{t_1}^{t_2} (\eta(t)-\xi(t), \varphi(dt)) = \lim_{n\to\infty} \int_{t_1}^{t_2} (\eta(t)-\xi_n(t), \varphi_n(dt)) \ge 0,$$

and the proof is finished.

We proceed to the existence problem for (2.5) assuming that $w \in \mathbf{C}(\mathbf{R}_+, \mathbf{R}^d)$. We begin with the special case when D satisfies the following condition.

CONDITION (A). There exist a unit vector $\mathfrak{e}$ and a constant $c>0$ such that $(\mathfrak{e}, \boldsymbol{n}) \geq c$ for any $\boldsymbol{n} \in \bigcup_{y \in \partial D} \mathcal{N}_y(D)$.

LEMMA 2.6. *Assume that D satisfies the condition (A). Then, there exists a solution ξ of* (2.5) *for any $w \in \mathbf{C}(\mathbf{R}_+, \mathbf{R}^d)$, and for $0 \leq s < t$*

$$|\xi(t) - \xi(s)| \leq K\varDelta_{s,t}, \tag{2.7a}$$

$$|\varphi|(t) - |\varphi|(s) \leq K'\varDelta_{s,t}, \tag{2.7b}$$

where K and K′ are constants depending only upon the constant c in the condition (A); $\varDelta_{s,t} = \sup_{s \leq t_1 < t_2 \leq t} |w(t_2) - w(t_1)|$.

PROOF. For each integer $n \geq 1$ we define $w_n \in \mathbf{D}(\mathbf{R}_+, \mathbf{R}^d)$ by $w_n(t) = w\left(\frac{k}{n}\right)$ for $\frac{k-1}{n} \leq t < \frac{k}{n}$, $k \geq 1$. Then w_n converges to w uniformly on compacts as $n \to \infty$, and by Lemma 2.1 there exists a solution ξ_n of $\xi_n = w_n + \varphi_n$. We put

$$\varDelta_{n,s,t} = \sup_{s \leq t_1 < t_2 \leq t} |w_n(t_2) - w_n(t_1)|,$$

$$K_{n,s,t} = |\varphi_n|(t) - |\varphi_n|(s),$$

and notice that

$$\begin{aligned}(\mathfrak{e}, \xi_n(t) - \xi_n(s)) &= (\mathfrak{e}, w_n(t) - w_n(s)) + (\mathfrak{e}, \varphi_n(t) - \varphi_n(s)) \\ &\geq (\mathfrak{e}, w_n(t) - w_n(s)) + cK_{n,s,t},\end{aligned}$$

that is,

$$K_{n,s,t} \leq \{|\xi_n(t) - \xi_n(s)| + \varDelta_{n,s,t}\}/c. \tag{2.8}$$

On the other hand, (2.6b) yields

$$\begin{aligned}|\xi_n(t) - \xi_n(s)|^2 &\leq \varDelta_{n,s,t}^2 + 2K_{n,s,t}\varDelta_{n,s,t} \\ &\leq \varDelta_{n,s,t}^2 + \varepsilon^2 K_{n,s,t}^2 + \varDelta_{n,s,t}^2/\varepsilon^2,\end{aligned}$$

that is,

$$|\xi_n(t) - \xi_n(s)| \leq \left(1 + \frac{1}{\varepsilon}\right)\varDelta_{n,s,t} + \varepsilon K_{n,s,t}, \qquad \varepsilon > 0.$$

This combined with (2.8) implies

$$|\xi_n(t) - \xi_n(s)| \leq \left(1 + \frac{1}{\varepsilon} + \frac{\varepsilon}{c}\right)\varDelta_{n,s,t} + \frac{\varepsilon}{c}|\xi_n(t) - \xi_n(s)|,$$

and therefore

$$(2.9) \qquad |\xi_n(t) - \xi_n(s)| \leq K\Delta_{n,s,t}, \quad K_{n,s,t} \leq K'\Delta_{n,s,t},$$

where K is the minimum of $\left(1+\frac{1}{\varepsilon}+\frac{\varepsilon}{c}\right)\left(1-\frac{\varepsilon}{c}\right)^{-1}$ as ε ranges over the interval $(0, c)$ and $K'=(1+K)/c$. In particular, $|\varphi_n|(T)$ $(=K_{n,0,T})$ is bounded and so by Lemma 2.5 ξ_n converges uniformly on compacts to the solution ξ of (2.5) as $n\to\infty$. This proves the existence part. The estimates (2.7a) and (2.7b) are also immediate from (2.9). The proof is finished.

Next, we introduce the condition (B) for a convex domain D.

CONDITION (B). There exist $\varepsilon>0$ and $\delta>0$ such that for any $x\in\partial D$ we can find an open ball $B_\varepsilon(x_0)=\{y\in\mathbf{R}^d: |y-x_0|<\varepsilon\}$ satisfying $B_\varepsilon(x_0)\subset D$ and $|x-x_0| \leq\delta$.

We can easily see that the condition (B) is always satisfied if D is bounded or if $d=2$. We now assume that D satisfies the condition (B) and for a point $x\in\partial D$ put

$$B(x) = \{y\in\mathbf{R}^d: |y-x| < \varepsilon/2\},$$

$$D_x = \bigcap_{y\in\partial D\cap\overline{B(x)}} \ \bigcap_{H\in\mathscr{H}_y(D)} H(D),$$

where $H(D)$ denotes the open half space bounded by a supporting hyperplane H and containing D. Then D_x is a convex domain satisfying the condition (A) with

$$\mathfrak{e} = (x_0 - x)/|x_0 - x|, \quad c = \varepsilon/2\delta.$$

THEOREM 2.1. (i) *Assume that D satisfies the condition (B). Then there exists a unique solution of* (2.5) *if $w\in\mathbf{C}(\mathbf{R}_+, \mathbf{R}^d)$, and the solution depends continuously upon w with respect to the compact uniform topology.* (ii) *Let D be a general convex domain and $\{w_n\}_{n\geq1}$ be a sequence in $\mathbf{C}(\mathbf{R}_+, \mathbf{R}^d)$ such that $\xi_n=w_n+\varphi_n$ has a solution for each n. Assume that w_n and ξ_n converge to w and ξ uniformly on compacts as $n\to\infty$, respectively. Then ξ is a solution of* (2.5).

PROOF. (i) If we put $T_0=\inf\{t\geq0: w(t)\notin\bar{D}\}$, then $\xi_0(t)\equiv w(t)$ $(0\leq t\leq T_0)$ is the solution of (2.5) for $0\leq t\leq T_0$. Assuming that the solution ξ_{n-1} of (2.5) is constructed for $0\leq t\leq T_{n-1}$ $(n\geq1)$, we now extend it beyond T_{n-1} as follows. Put $w^{(n)}(t)=w(T_{n-1}+t)$, $t\geq0$, let $\xi^{(n)}$ be the solution of $\xi^{(n)}=w^{(n)}+\varphi^{(n)}$ on $\bar{D}_{\xi_{n-1}(T_{n-1})}$ and again put

$$t_n = \inf\{t \geq T_{n-1}: |\xi^{(n)}(t - T_{n-1}) - \xi^{(n)}(0)| = \varepsilon/2\},$$

$$T_n = \inf\{t \geq t_n: \xi^{(n)}(t_n - T_{n-1}) + w(t) - w(t_n) \notin \bar{D}\},$$

$$\xi_n(t) = \begin{cases} \xi_{n-1}(t) & , \quad 0 \le t \le T_{n-1}, \\ \xi^{(n)}(t - T_{n-1}) & , \quad T_{n-1} \le t \le t_n, \\ \xi^{(n)}(t_n - T_{n-1}) + w(t) - w(t_n), & \quad t_n \le t \le T_n. \end{cases}$$

Then ξ_n is the solution of (2.5) on $\bar{D}$ for $0 \le t \le T_n$. Repeating this argument, we obtain an increasing sequence $\{T_n\}$ and a continuous function ξ defined for $t < T_\infty = \lim T_n$ such that ξ is the solution of (2.5) for $0 \le t < T_\infty$. The associated function φ is flat on each interval (t_n, T_n) and (2.7) holds for $s, t \in [T_{n-1}, t_n]$. Therefore, (2.7) holds for $s, t \in [T_{n-1}, T_n]$ with constants K and K' depending only on $c = \varepsilon/2\delta$. But, (2.7a) with $s = T_{n-1}$ and $t = t_n$ implies

$$\varepsilon/2K \le \Delta_{T_{n-1}, t_n} \le \Delta_{T_{n-1}, T_n},$$

from which we can claim as follows: If $h > 0$ is so small that $\Delta_T(h) < \varepsilon/2K$ where

$$\Delta_T(h) = \max\{|w(t) - w(s)| : 0 \le s, t \le T \quad \text{and} \quad |t - s| \le h\},$$

T being an arbitrarily fixed positive constant, then $T_n \le T$ implies $T_n - T_{n-1} > h$. In other words, $T_n > T$ for all $n > T/h$. This fact implies the followings:

(2.10) $T_n = \infty$, that is, there exists a solution of (2.5) for $0 \le t < \infty$.

(2.11) For $0 \le s < t \le T$ the solution satisfies

(a) $|\xi(t) - \xi(s)| \le \left(\frac{T}{h} + 1\right) K \Delta_{s,t}$,

(b) $|\varphi|(t) - |\varphi|(s) \le \left(\frac{T}{h} + 1\right) K' \Delta_{s,t}$.

Next, let $\{w_n\}_{n \ge 1}$ be a sequence in $\mathbf{C}(\mathbf{R}_+, \mathbf{R}^d)$ converging to w uniformly on compacts and let $\xi_n = w_n + \varphi_n$. Then (b) of (2.11) applied to φ_n yields

$$|\varphi_n|(t) - |\varphi_n|(s) \le \left(\frac{T}{h_n} + 1\right) K' \Delta_{n,s,t}, \qquad 0 \le s < t \le T.$$

Here h_n depends upon w_n, but it can be chosen to be independent of w_n for all sufficiently large n because $w_n \to w$. Therefore the above inequality on $|\varphi_n|$ implies that $\{|\varphi_n|(T)\}$ is bounded and hence by Lemma 2.5 $\xi_n \to \xi$ uniformly on $[0, T]$.

(ii) Let $T > 0$ be any fixed constant. Then there exists N such that

$$\sup_n \max_{0 \le t \le T} |\xi_n(t)| < N.$$

For such an N both ξ_n and ξ are the solutions of $\xi_n = w_n + \varphi_n$ and $\xi = w + \varphi$, $0 \le t \le T$, for the domain $D_N = D \cap \{|x| < N\}$. Since D_N satisfies the condition (B), we can apply the result of (i) to conclude that ξ is the solution of (2.5) $(0 \le t \le T)$ for D_N and hence for D. The proof is finished.

REMARK 2.3. Even if D does not satisfy the condition (B) (so $d \ge 3$ and D is unbounded), for each $x \in \partial D$ we can find an open ball $B_\varepsilon(x_0) = \{y \in \mathbf{R}^d : |y - x_0| < \varepsilon\}$ inside D (but now ε or $|x - x_0|^{-1}$ is not bounded away from zero as x moves on ∂D). Therefore, in a manner similar to the proof of Theorem 2.1, (i), we can construct the solution of (2.5) for $t < T_\infty = \lim T_n$. I can neither prove that $T_\infty = \infty$ in general nor give an example in which $T_\infty < \infty$.

§3. A stochastic version of (2.5)

The purpose of this section is to remove the condition (B) in the existence of global solutions of (2.5) by taking w from sample paths of a continuous semi-martingale.

Let $(\Omega, \mathscr{F}, P)$ be a complete probability space with an increasing family $\{\mathscr{F}_t\}_{t \ge 0}$ of sub-σ-fields of $\mathscr{F}$; it is assumed that each $\mathscr{F}_t$ contains all P-negligible sets and $\mathscr{F}_t = \bigcap_{\varepsilon > 0} \mathscr{F}_{t+\varepsilon}$. Let D be a convex domain as before.

THEOREM 3.1. *Let $\{M(t)\}$ be an $\mathbf{R}^d$-valued process with $M(0) \in \bar{D}$ such that each component is a continuous local $\mathscr{F}_t$-martingale and $\{A(t)\}$ be an $\mathbf{R}^d$-valued, continuous and $\mathscr{F}_t$-adapted process of bounded variation with $A(0) = 0$. Then there exists a unique $\mathscr{F}_t$-adapted solution $\{X(t)\}$ of*

$$X(t) = M(t) + A(t) + \Phi(t). \tag{3.1}$$

Moreover, for $f \in \mathbf{C}^2(\mathbf{R})$ with $f' \ge 0$ on $\mathbf{R}_+$ and $0 \le s \le t$ we have

$$\begin{aligned} &f(|X(t) - X(s)|^2) \\ &\le f(0) + 2\int_s^t f' \sum_i (X^i(\tau) - X^i(s))(M^i(d\tau) + A^i(d\tau)) \\ &\quad + 2\int_s^t f'' \sum_{i,j} (X^i(\tau) - X^i(s))(X^j(\tau) - X^j(s)) d[M^i, M^j] \\ &\quad + \int_s^t f' \sum_i d[M^i, M^i], \end{aligned} \tag{3.2}$$

where f', f'' are evaluated at $|X(\tau) - X(s)|^2$ and $[M^i, M^j]$ denotes the quadratic variation process.

REMARK 3.1. By a solution of (3.1), we mean a $\bar{D}$-valued process $\{X(t)\}$ which satisfies (3.1) almost surely, under the condition that almost all sample paths of $\{\Phi(t)\}$ are associated with those of $\{X(t)\}$.

PROOF. Let $\tau(t)$ be the inverse function of

$$\theta(t) = t + \sum_{i=1}^{d} [M^i, M^i] + |A|(t),$$

and put

$$\mathscr{F}_t^* = \mathscr{F}_{\tau(t)}, \; M^*(t) = M(\tau(t)), \quad A^*(t) = A(\tau(t)).$$

Then $\{M^*(t)\}$ is a continuous $\mathscr{F}_t^*$-martingale and $\{A^*(t)\}$ is a continuous $\mathscr{F}_t^*$-adapted process of bounded variation, satisfying

$$0 \le \sum_{i,j=1}^{d} x^i x^j d[M^{*i}, M^{*j}] \le |x|^2 dt, \qquad x \in \mathbf{R}^d, \tag{3.3a}$$

$$A^*(t) = \int_0^t a^*(s)ds, \qquad |a^*(t)| \le 1. \tag{3.3b}$$

Moreover, once we obtain the $\mathscr{F}_t^*$-adapted solution of $X^*(t) = M^*(t) + A^*(t) + \Phi^*(t)$, the $\mathscr{F}_t$-adapted solution of (3.1) can be obtained by $X(t) = X^*(\theta(t))$. Therefore, in proving the theorem we may assume that $\{M(t)\}$ and $\{A(t)\}$ themselves satisfy (3.3) (without the symbol $*$).

First we prove (3.2) assuming the existence of the $\mathscr{F}_t$-adapted solution $\{X(t)\}$. An application of Itô's formula yields

$$f(|X(t) - X(s)|^2) = \text{the right hand side of (3.2)}$$
$$+ \int_s^t f' \cdot (X(\tau) - X(s), \Phi(d\tau)).$$

But the last term is non-positive by (2.4′).

In order to prove the existence of the solution, we first consider the equation for $D_n = D \cap \{|x| < n\}$. Taking a point x^* in D_n (we may consider only those n for which $D_n \neq \emptyset$), we put

$$M_n(t) = \mathbf{1}_{D_n}(M(0))M(t) + \mathbf{1}_{\mathbf{R}^d - D_n}(M(0))x^*,$$

$$A_n(t) = \mathbf{1}_{D_n}(M(0))A(t).$$

Since D_n is bounded and hence satisfies the condition (B), by Theorem 2.1 there exists a unique $\mathscr{F}_t$-adapted solution $\{X_n(t)\}$ of $X_n(t) = M_n(t) + A_n(t) + \Phi_n(t)$ for D_n. If we put

$$T_n = \inf\{t \ge 0 : |X_n(t)| = n\},$$

then $\{X_n(t \wedge T_n)\}$ is again the solution of $X_n(t \wedge T_n) = M_n(t \wedge T_n) + A_n(t \wedge T_n) + \Phi_n(t \wedge T_n)$ for D_n, and so (3.2) can be applied to $|X_n(t \wedge T_n) - X_n(0)|^2$. Thus, by taking the expectation we have

$$E\{|X_n(t \wedge T_n) - X_n(0)|^2\} \le 2\int_0^t E\{|X_n(s \wedge T_n) - X_n(0)|\}ds + t$$

$$\le \int_0^t E\{|X_n(s \wedge T_n) - X_n(0)|^2\}ds + 2t,$$

and hence

$$E\{|X_n(t \wedge T_n) - X_n(0)|^2\} \le 2te^t,$$

that is, $E\{|X_n(t \wedge T_n)|^2\}$ is bounded in n for each fixed t. Therefore, for each t

$$P\{T_n \le t\} \le E\{|X_n(t \wedge T_n)|^2\}/n^2 \to 0 \quad \text{as} \quad n \to \infty. \tag{3.4}$$

On the other hand, the uniqueness lemma in § 2 implies that

$$T_n \le T_m \quad \text{and} \quad X_n(t) = X_m(t) \qquad \text{for} \quad t \le T_n$$

hold on the set $\{M(0) \in \bar{D}_n\}$ if $n<m$. This fact combined with (3.4) enables us to define $\{X(t)\}$ almost surely by

$$X(t) = X_n(t) \quad \text{on} \quad \{M(0) \in \bar{D}_n\} \cap \{t \le T_n\}.$$

Thus defined $\{X(t)\}$ is clearly the $\mathscr{F}_t$-adapted solution of (3.1).

§ 4. Stochastic differential equation with reflection

Let D be a convex domain in $\mathbf{R}^d$ and $\{\Omega, \mathscr{F}, P; \mathscr{F}_t\}$ satisfy the same condition as in § 3. We suppose that an $\mathscr{F}_t$-adapted r-dimensional Brownian motion $B(t)=(B^1(t),\dots, B^r(t))$ with $B(0)=0$ is given; that is, $\{B(t)\}$ is an $\mathscr{F}_t$-adapted continuous process and for $0 \le s \le t$, $\xi \in \mathbf{R}^d$

$$E\{e^{i(\xi, B(t)-B(s))}|\mathscr{F}_s\} = e^{-(t-s)|\xi|^2/2}, \quad \text{a.s.}$$

Given an $\mathbf{R}^d \otimes \mathbf{R}^r$-valued function $\sigma(t, x)=\{\sigma_k^i(t, x)\}$ and an $\mathbf{R}^d$-valued function $b(t, x)=\{b^i(t, x)\}$, both being defined on $\mathbf{R}_+ \times \bar{D}$, we consider the stochastic differential equation with reflection

$$dX = \sigma(t, X)dB + b(t, X)dt + d\Phi, \quad X(0) = x, \tag{4.1}$$

or equivalently

$$X^i(t) = x^i + \sum_{k=1}^r \int_0^t \sigma_k^i(s, X(s))dB^k(s) + \int_0^t b^i(s, X(s))ds + \Phi^i(t), \tag{4.1'}$$

where $x=(x^1,\dots, x^d) \in \bar{D}$. Our problem is to find an $\mathscr{F}_t$-adapted $\bar{D}$-process $\{X(t)\}$ under the condition that $\{\Phi(t)\}$ is an associated process of $\{X(t)\}$. It is always assumed that $\sigma(t, x)$ and $b(t, x)$ are Borel measurable in (t, x).

THEOREM 4.1. *If there exists a constant $K>0$ such that*

$$\|\sigma(t, x) - \sigma(t, y)\| \leq K|x - y|, \ |b(t, x) - b(t, y)| \leq K|x - y|, \tag{4.2}$$

$$\|\sigma(t, x)\| \leq K(1 + |x|^2)^{1/2}, \ \|b(t, x)\| \leq K(1 + |x|^2)^{1/2}, \tag{4.3}$$

then there exists a (pathwise) unique $\mathscr{F}_t$-adapted solution of (4.1) *for any $x \in \bar{D}$.*

PROOF. First we prove the pathwise uniqueness of the solution. Let $\{X(t)\}$ and $\{Y(t)\}$ be $\mathscr{F}_t$-adapted solutions of (4.1). Then by the first inequality in Remark 2.2

$$\begin{aligned} |X(t) - Y(t)|^2 \leq & \left|\int_0^t \sigma(s, X)dB - \int_0^t \sigma(s, Y)dB\right|^2 \\ & + 2\int_0^t (X(s) - Y(s), b(s, X) - b(s, Y))ds \\ & + \text{the remainder.} \end{aligned}$$

Writing the remainder term explicitly, we can see that it has zero expectation[1] and hence

$$\begin{aligned} E\{|X(t) - Y(t)|^2\} \leq & E\int_0^t \|\sigma(s, X) - \sigma(s, Y)\|^2 ds \\ & + E\int_0^t |X(s) - Y(s)|^2 ds + E\int_0^t |b(s, X) - b(s, Y)|^2 ds \\ \leq & (2K^2 + 1)\int_0^t E\{|X(s) - Y(s)|^2\}ds. \end{aligned} \tag{4.4}$$

Therefore, $E\{|X(t) - Y(t)|^2\} = 0$.

We give the existence proof, first assuming that D is bounded. By Theorem 2.1, we can define a sequence $\{X^{(n)}(t)\}$ of $\bar{D}$-processes by

$$X^{(0)}(t) = x$$

$$X^{(n)}(t) = x + \int_0^t \sigma(s, X^{(n-1)})dB + \int_0^t b(s, X^{(n-1)})ds + \Phi^{(n)}, \ n \geq 1.$$

Then, as in (4.4) we have

1) Because we do not know beforehand that $|X(t) - Y(t)|^2$ and the remainder term are integrable, we must employ the following truncation argument. Let T_n be the infimum of $t \geq 0$ at which

$$\int_0^t \left|X(s) - Y(s)\right|^2 ds + |\Phi|(t) + |\Psi|(t) = n,$$

and derive (4.4) for $X(\cdot \wedge T_n)$ and $Y(\cdot \wedge T_n)$; it then follows that $E\{|X(t\wedge T_n) - Y(t\wedge T_n)|^2\} = 0$; now let $n \uparrow \infty$.

$$E\{|X^{(n+1)}(t)-X^{(n)}(t)|^2\}\leq(2K^2+1)\int_0^t E\{|X^{(n)}(s)-X^{(n-1)}(s)|^2\}ds.$$

Therefore, by a routine argument we see that

$$\int_0^t \sigma(s, X^{(n)})dB+\int_0^t b(s, X^{(n)})ds$$

is convergent uniformly on compacts as $n\to\infty$ (a.s.). Consequently, from Theorem 2.1, (i), it follows that $\{X^{(n)}(t)\}$ is also convergent uniformly on compacts as $n\to\infty$ (a.s.) and the limit process $\{X(t)\}$ is an $\mathscr{F}_t$-adapted solution of (4.1). When D is unbounded, we put $D_n=D\cap\{|x|<n\}$ and consider the solution $\{X_n(t)\}$ of (4.1) on $\bar{D}_n$. Let T_n be the infimum of $t\geq 0$ at which $|X_n(t)|=n$. Then, applying (3.2) to $|X_n(t\wedge T_n)-x|^2$ and taking the expectation, we have

$$E\{|X_n(t\wedge T_n-x|^2)\leq 2E\int_0^{t\wedge T_n}(X_n(s)-x, b(s, X_n))ds$$
$$+E\int_0^{t\wedge T_n}\sum_{i,k}\sigma_k^i(s, X_n)^2 ds,$$

and hence by making use of (4.3)

$$E\{|X_n(t\wedge T_n)-x|^2\}\leq\int_0^t E\{|X_n(s\wedge T_n)-x|^2\}ds$$
$$+2K^2\int_0^t E\{1+|X_n(s\wedge T_n)|^2\}ds.$$

Therefore, by Gronwall's inequality we see that $E\{|X_n(t\wedge T_n)|^2\}$ is bounded in n for each fixed t and hence $P\{T_n\leq t\}\leq E\{|X_n(t\wedge T_n)|^2\}/n^2\to 0$ as $n\to\infty$. On the other hand, by the uniqueness already proved we have $T_n\leq T_m$ and $X_n(t)=X_m(t)$ for $t\leq T_m$ (a.s.) provided $n<m$, and therefore we can construct an $\mathscr{F}_t$-adapted solution $\{X(t)\}$ of (4.1) for D by $X(t)=X_n(t)$, $0\leq t\leq T_n$. This completes the proof.

THEOREM 4.2. *If $\sigma(t, x)$ and $b(t, x)$ are bounded continuous on $\mathbf{R}_+\times\bar{D}$, then on some probability space $(\Omega, \mathscr{F}, P)$ we can find an r-dimensional Brownian motion $\{B(t)\}$ in such a way that* (4.1) *has a solution.*

PROOF. We choose sequences $\{\sigma_n(t, x)\}$ and $\{b_n(t, x)\}$ such that (i) $\sigma_n\to\sigma$ and $b_n\to b$ boundedly and uniformly on compacts as $n\to\infty$, and (ii) σ_n and b_n satisfy the Lipschitz condition. Then there exists a solution $\{X_n(t)\}$ of (4.1) with coefficients σ_n and b_n, for each n. Making use of (3.2) with $f(u)=u^p$, $p\geq 1$,

$$E\{|X_n(t)-X_n(s)|^{2p}\}\leq c_p\Big\{E\int_s^t|X_n(\tau)-X_n(s)|^{2p-1}d\tau$$
$$+E\int_s^t|X_n(\tau)-X_n(s)|^{2p-2}d\tau\Big\},$$

where c_p is some constant depending on p but not on n. Therefore $E\{|X_n(t) - X_n(s)|^4\} \leq c|t-s|^2$ for $0 \leq s, t \leq T$, where c is some constant depending on T but not on n, T being an arbitrary positive constant. This implies that the family of probability measures on $\mathbf{C}(\mathbf{R}_+, \bar{D}) \times \mathbf{C}(\mathbf{R}_+, \mathbf{R}^d)$ induced by $\{(X_n(t), B(t))\}$ is tight. Therefore, noticing the result (ii) of Theorem 2.1 we can fill the rest of our proof exactly in the same way as in the case of the whole space $\mathbf{R}^d$ (see [5]). The proof is finished.

The following theorem, due to O. Tanbara (unpublished), can be proved in the same way as in [9] making use of the estimate (3.2).

THEOREM 4.3. *Let* ρ *and* $\bar{\rho}$ *satisfy*

$$\int_{0+} \{\rho^2(u)u^{-1} + \bar{\rho}(u)\}^{-1} du = \infty, \tag{4.5}$$

$$\rho^2(u)u^{-1} + \bar{\rho}(u) \quad \text{is concave.} \tag{4.6}$$

Then, for any $\sigma(t, x)$ *and* $b(t, x)$ *satisfying*

$$\|\sigma(t, x) - \sigma(t, y)\| \leq \rho(|x-y|), \quad |b(t, x) - b(t, y)| \leq \bar{\rho}(|x-y|),$$

the pathwise uniqueness of solutions holds for (4.1).

References

[1] N. Ikeda, On the construction of two-dimensional diffusion processes satisfying Wentzell's boundary conditions and its application to boundary value problems, Mem. Coll. Sci. Univ. Kyôto, **33** (1961), 367–427.

[2] H. P. McKean, A. Skorohod's integral equation for a reflecting barrier diffusion, J. Math. Kyôto Univ., **3** (1963), 86–88.

[3] H. P. McKean, Stochastic Integrals, Academic Press, 1969.

[4] A. V. Skorohod, Stochastic equations for diffusion processes in a bounded region 1, 2, Theor. Veroyatnost. i Primenen. **6** (1961), 264–274; **7** (1962), 3–23.

[5] A. V. Skorohod, Studies in the Theory of Random Processes, Addison Wesley, 1965.

[6] D. W. Stroock and S. R. S. Varadhan, Diffusion processes with boundary conditions, Comm. Pure Appl. Math., **24** (1971), 147–225.

[7] M. Tsuchiya, On the stochastic differential equation for a two-dimensional Brownian motion with boundary condition, to appear.

[8] S. Watanabe, On stochastic differential equations for multi-dimensional diffusion processes with boundary conditions, J. Math Kyôto Univ., **11** (1971), 169–180.

[9] S. Watanabe and T. Yamada, On the uniqueness of solutions of stochastic differential equations II, J. Math. Kyôto Univ., **11** (1971), 553–563.

Department of Mathematics,
Faculty of Science,
Hiroshima University

SOME PROBABILISTIC PROBLEMS IN THE SPATIALLY HOMOGENEOUS BOLTZMANN EQUATION

Hiroshi Tanaka
Department of Mathematics
Faculty of Science and Technology
Keio University, Yokohama, Japan

1. Introduction

1.1. The master equation approach to the spatially homogeneous Boltzmann equation initiated by Kac [1] and continued by McKean [3], Grünbaum [5] and others is briefly reviewed together with some new results on

(I) propagation of chaos (law of large numbers),
(II) fluctuation (central limit theorem).

We are interested in the time evolution of the velocities $X_1^{(n)}(t), \dots, X_n^{(n)}(t)$ of n particles moving in the space R^3 under certain binary interaction (collision). We assume that $(X_1^{(n)}(t), \dots, X_n^{(n)}(t))$ is a Markov process on R^{3n} with generator

$$(K_n\varphi)(x_1,\dots,x_n)$$
$$= \frac{1}{n}\sum_{1\le i<j\le n}\int_{(0,\pi)\times(0,2\pi)}\left\{\varphi(\dots,x_i',\dots,x_j',\dots)-\varphi(x_1,\dots,x_n)\right\}Q(x_i,x_j,\theta)\,d\theta\, d\epsilon ,$$

where x_i' and x_j' are the post-collision velocities of the i-th and j-th particles with velocities x_i and x_j. If $S(x,y)$ denotes the 2-dimensional sphere with center $(x+y)/2$ and diameter $|x-y|$, the x_i' and x_j' are always on $S(x_i,x_j)$ or more precisely $S(x_i',x_j')=S(x_i,x_j)$. We now take a spherical coordinate system on $S(x_i,x_j)$ with north pole x_i and denote by θ (resp. ϵ) the colatitude (resp. longitude) of x_i'. The function $Q(x,y,\theta)$, which is determined by the interparticle (binary) potential, is assumed to be nonnegative and depend only on $|x-y|$, $x+y$ and θ. The forward equation

$$\frac{d}{dt}\langle u(t),\varphi\rangle = \langle u(t),K_n\varphi\rangle, \quad \varphi = \text{(test) function on } R^{3n} \tag{1}$$

describing the Markov process $X^{(n)}(t)$ is called the n-particle <u>master equation</u> corresponding to the following spatially homogeneous <u>Boltzmann equation</u>

(2) $$\frac{\partial u}{\partial t} = \int_{(0,\pi)\times(0,2\pi)\times R^3} (u'u_1' - uu_1)Q(x,x_1,\theta)d\theta d\epsilon\, dx_1 ,$$

where $u = u(t,x)$, $u_1=u(t,x_1)$, $u'=u(t,x')$, $u_1'=u(t,x_1')$; x' and x_1' are the post-collision velocities, i.e., $(x,x_1) \to (x',x_1')$ by "collision"; the notation $\langle u,\varphi\rangle$ in general stands for the integral of a function φ with respect to a measure u. As in [7] we consider a weak version of (2);

(3) $$\frac{d}{dt}\langle u(t),\varphi\rangle = \langle u(t)\otimes u(t),K\varphi\rangle ,$$

or equivalently

(3#) $$\frac{d}{dt}\langle u(t),\varphi\rangle = \langle u(t)\otimes u(t),K^{\#}\varphi\rangle ,$$

where φ is a (test) function R^3 and

(4) $$(K\varphi)(x,x_1) = \int_{(0,\pi)\times(0,2\pi)} \{\varphi(x') - \varphi(x)\}Q(x,x_1,\theta)d\theta d\epsilon ,$$

$$(K^{\#}\varphi)(x,x_1) = \{(K\varphi)(x,x_1) + (K\varphi)(x_1,x)\}/2 .$$

By a weak solution of (2) we mean a probability measure solution of (3) $(=(3^{\#}))$.

1.2. <u>Propagation of chaos</u>. Let $\{u_n, n>1\}$ be a sequence of probability measures, each u_n being a symmetric probability measure on the n-fold product space $R^{3n} = R^3\times\ldots\times R^3$. Let u be a probability measure on R^3 . Then a sequence $\{u_n\}$ is said to be u-chaotic if

$\langle u_n,\varphi_1\otimes\cdots\otimes\varphi_m\otimes 1\otimes\cdots\otimes 1\rangle \to \prod_{k=1}^{m}\langle u,\varphi_k\rangle$ as $n\to\infty$ for any $\varphi_1,\ldots,$ $\varphi_m \in C_0(R^3)$, $m\geq 1$. The relation between the master equation (1) and the corresponding Boltzmann equation (2) was made clear by Kac [I] through the following propagation of chaos : <u>Let</u> $u_n(t)$ <u>be the solution of</u> (1) <u>with initial value</u> u_n <u>and assume that</u> $\{u_n, n>1\}$ <u>is</u> u-<u>chaotic</u>. <u>Then</u> $\{u_n(t), n>1\}$ <u>is also</u> u(t)-<u>chaotic where</u> $u(t)$ <u>is the</u> (<u>weak</u>) <u>solution</u> <u>of</u> (2) <u>with initial value</u> u . We next introduce the normalized occupation number

$$\overline{X}_n(t) = \frac{1}{n}\sum_{k=1}^{n}\delta_{X_k^{(n)}(t)}$$

where δ_x denotes the δ-distribution at x . We can easily show that

$\{u_n\}$ is u-chaotic if and only if the law of large numbers holds for $\{u_n\}$ in the following sense:

$$\frac{1}{n}\sum_{k=1}^{n} \delta_{X_k^{(n)}} \longrightarrow u \ , \quad n \to \infty \text{ (in probability)}$$

where $X^{(n)} = (X_1^{(n)},\ldots,X_n^{(n)})$ is a u_n-distributed random vector. Therefore, since $X^{(n)}(t)$ is $u_n(t)$-distributed provided that $X^{(n)}(0)$ is u_n-distributed, the propagation of chaos is equivalent to the following <u>law of large numbers</u>:

$$(5) \quad \overline{X}_n(0) \to u, \quad n \to \infty \text{ (in prob.)} \Rightarrow \overline{X}_n(t) \to u(t), \quad n \to \infty \text{ (in prob.)}$$

where $u(t)$ is the same as before.

The proof of the propagation of chaos was first given by Kac [1] for the case of 1-dimensional Kac's model of Maxwellian molecules; McKean [3] gave a beautiful proof for some special case including cutoff model of the Boltzmann equation of Maxwellian molecules. Murata [6] treated the 2-dimensional non-cutoff Maxwellian model. Grünbaum [5] discussed the case of a considerably wide (cutoff) class; his discussions covered the gas of hard spheres but under some assumption which was unverified though very believable. In §2 of this article the propagation of chaos will be proved in the following two cases.

(i) $\displaystyle\int_0^{\pi} Q(x,y,\theta)d\theta \le \text{const.}(1+|x|^2+|y|^2)$.

(ii) (Maxwellian type) $Q(x,y,\theta)$ is a function $Q(\theta)$ of θ alone satisfying $\displaystyle\int_0^{\pi} \theta^2 Q(\theta)d\theta < \infty$.

It is to be noted that the case (i) includes the gas of hard spheres ($Q(x,y,\theta) = \text{const.}|x-y|\sin\theta$) while the case (ii) includes the 3-dimensional Maxwellian molecules with the inverse 5-th power interparticle repulsive force (in such a case $Q(\theta) \sim \text{const.}\theta^{-3/2}, \theta \downarrow 0$) .

1.3. <u>Fluctuation</u>. Since the normalized occupation number $\overline{X}_n(t)$ is fluctuating about a solution $u(t)$ of the Boltzmann equation (2), the next problem is to study the asymptotic behavior of

(6) $$Y_n(t) = \sqrt{n}(\overline{X}_n(t) - u(t))$$

as $n \to \infty$. The case of McKean's 2-state model of Maxwellian moleclues was first discussed by Kac [2] and then by McKean [4] in detail. As found in heuristic discussion by [4] for the case of gas of hard spheres, the limit process of $Y_n(t)$ must be, in general, an infinite dimensional Ornstein-Uhlenbeck process. Rigorous discussions in the case of Kac's 1-dimensional model of Maxwellian molecules were given by Tanaka [7] (equilibrium case) from the point of view of a limit theorem on Markov processes taking values of tempered distributions. Non-equilibrium case was then treated by Uchiyama [8]. Recently, Uchiyama [9] discussed the equilibrium case of cutoff type including gas of hard spheres. In §3 of this article the fluctuation theory (=central limit theorem) will be discussed in the case (ii) (Maxwellian type) along the same lines as in [7]. The emphasis here is on the non-cutoff property.

Fundamentally our method is to derive appropriate convergences of Markov processes $\overline{X}_n(t)$ and $Y_n(t)$ knowing the convergence of their generators (a martingale problem approach will then be useful), except for the treatment of chaos propagation in the case (ii) where the cutoff approximation will be used. Proofs are only outlined; details will appear elsewhere.

2. Propagation of chaos

2.1. Case (i). Let $\xi_0=1$, $\xi_1=x$ and $\xi_k=|x|^k$ for $k\geq 2$ $(x\in R^3)$.

Theorem 1. The function $Q(x,y,\theta)$, depending only on $|x-y|$, $x+y$ and θ , is assumed to satisfy the condition (i) of §1. In addition, we assume that

(7) $K\varphi$ is continuous provided that φ is bounded and continuous.

Let u_n be the initial distribution of $X^{(n)}(\cdot)$ and assume that $\{u_n\}$ is a u-chaotic sequence satisfying

$$\langle u,\xi_6\rangle<\infty, \quad \sup_{n\geq 1}\langle u_n, \xi_6\otimes 1\otimes\cdots\otimes 1\rangle<\infty .$$

Then for any $\varepsilon>0$ and T $(0<T<\infty)$

(8) $$P\left\{\sup_{0\leq t\leq T} \rho(\overline{X}_n(t), u(t))>\varepsilon\right\}\to 0$$

as $n\to\infty$, where ρ is any metric on $\mathcal{M}$ (the space of probability measures on R^3) which gives the usual vague topology on $\mathcal{M}$.

Remark. $u(t)$ in (8) is the solution of (3) (with $u(0)=u$) whose existence and uniquencess are guaranteed by the following Arkeryd's result [10]: Under the same assumption on $Q(x,y,\theta)$ as in Theorem 1 except for (7) (which is unnecessary here), for any initial value $u(o) = u$ satisfying $\langle u, \xi_4\rangle<\infty$ there exists a unique solution $u(t)$ of (3) such that $\langle u(t), \xi_4\rangle$ is bounded on each finite t-interval. The solution $u(t)$ also satisfies $\langle u(t), \xi_k\rangle = \langle u, \xi_k\rangle$, $k = 0,1,2$.

The proof of the theorem is sketched here. Choose a sequence $\{f_k, k\geq 1\}$ in $C_0(R^3)$ such that $\|f_k\|_\infty \leq 1$ and the set of all (finite) linear combinations of f_k's is dense in $C_0(R^3)$, and set $\rho(u, \tilde{u}) = \sum_{k=1}^{\infty} 2^{-k}|\langle u-\tilde{u}, f_k\rangle|$ for $u, \tilde{u}\in\mathcal{M}$. We first consider the special case in which $u_n(\mathcal{X}_n) = 1$, $n>1$, where $\mathcal{X}_n$ is the ball in R^{3n} of radius cn , the constant c being independent of n . Let $\mathcal{M}_0$ be the set of $u\in\mathcal{M}$ such that $\langle u, \xi_2\rangle \leq c$ and W be the space (endowed with the Skorohod topology) of $\mathcal{M}_0$-valued right continuous paths with left limits. Then $\overline{X}_n(t)$ is regarded as a Markov process with sample path in W . Denote by P_n the probability measure on W induced by the process $\overline{X}_n(t)$. We can prove the following lemmas (1^o-3^o) in which T is an arbitrary positive constant.

1^o. $E_n\{\rho(w(t_1),w(t_2))^2\rho(w(t_2),w(t_3))^2\}\leq \text{const.}|t_3-t_1|^2$ for $0\leq t_1<t_2<t_3\leq T$, where const. may depend on T . Therefore, $\{P_n\}$ is tight.

2^o. $\sup_n P_n\{\sup_{0\leq t\leq T}\langle w(t), \xi_4\rangle > N\} \longrightarrow 0$ as $N \to\infty$.

3^o. If P_∞ is any limit point of $\{P_n\}$, then

a) $\sup_{0\leq t\leq T}\langle w(t), \xi_4\rangle<\infty$, P_∞-a.s.,

b) $\langle w(t),\varphi\rangle - \langle w(0),\varphi\rangle - \int_0^t \langle w(s)\otimes w(s), K^{\#}\varphi\rangle ds=0$, P_∞-as, $\varphi\in C_0(R^3)$.

Since $\overline{X}_n(0) \to u$ in probability, $w(0)=u$ (P_∞-a.s.). Therefore, Arkeryd's uniqueness result applied to 3^o implies that $w(t)=u(t)$, $t\geq 0$, P_∞-a.s., i.e., $P_\infty = \delta_{u(\cdot)}$. This implies (8).

General case can be reduced to the spacial case by noticing that $u_n(\mathcal{X}_n) \to 1$ as $n\to\infty$ where $\mathcal{X}_n$ is defined as before choosing a constant $c>\langle u, \xi_2\rangle$.

2.2. Case (ii). We deal with this case by approximating $Q(x,y,\theta)$ by cutoff one. Let $Q_\varepsilon(\theta)=Q(\theta)$ (for $\varepsilon<\theta<\pi$) and $=0$ (for $0<\theta\leq\varepsilon$), and define $K^{(\varepsilon)}$ with $Q(\theta)$ replaced by $Q_\varepsilon(\theta)$ (cutoff) in the definition (4) of K . Let $u_\varepsilon(t)$ be the solution of

$$(3\varepsilon)\qquad \frac{d}{dt}\langle u(t),\varphi\rangle = \langle u(t)\otimes u(t), K^{(\varepsilon)}\varphi\rangle$$

with initial value u .

Theorem 2. (i) If $\langle u,\xi_2\rangle<\infty$, then $u_\varepsilon(t)$ converges weakly to some $u(t)$ (as $\varepsilon\downarrow 0$) which is a solution of (3) with initial value u . (ii) If $\langle u,\xi_2\rangle<\infty$ and if $\{u_n\}$ is a u-chaotic sequence satisfying

$$\lim_{n\to\infty}\langle u_n,\xi_2\otimes 1\otimes\ldots\otimes 1\rangle = \langle u,\xi_2\rangle,$$

then $\{u_n(t)\}$ is also u(t)-chaotic and

$$\lim_{n\to\infty}\langle u_n(t),\xi_2\otimes 1\otimes\ldots\otimes 1\rangle = \langle u(t),\xi_2\rangle,$$

where u(t) is the solution of (3) constructed in (i).

3. Fluctuation

We consider only the case of non-cutoff Maxwellian type. Thus we assume that a function $Q(\theta)$ satisfying (ii) of §1 is given. We set

$$g(x)=(2\pi)^{-3/2}e^{-|x|^2/2}, \quad g=g(x)dx, \quad e_0(x)=\sqrt{g(x)} .$$

We also assume that the initial distribution of $X^{(n)}(\cdot)$ is the n-fold product $g^{n\otimes}$. In this case we have $u_n(t)=g^{n\otimes}$ and $u(t)=g$, and (for notational convention) it is better to modify the definition (6) of $Y_n(t)$ as follows.

$$Y_n(t) = \sqrt{n}\{\overline{X}_n(t) - g\}/e_0(\cdot),$$

i.e.,
$$\langle Y_n(t),\varphi\rangle = n^{-1/2}\sum_{k=1}^{n}\left\{\frac{\varphi(X_k^{(n)}(t))}{e_0(X_k^{(n)}(t))} - \langle e_0,\varphi\rangle\right\}.$$

The fluctuation theory is to study the asymptotic behavior of $Y_n(t)$ as $n\to\infty$. Note that $Y_n(t)$ is a Markov process on the state space

$$\left\{\eta\in\mathcal{S}'(R^3):\eta=\sqrt{n}(\overline{x}-g)/e_0(\cdot), \quad \overline{x}=\frac{1}{n}\sum_{1}^{n}\delta_{x_k}, x_1,\ldots,x_n\in R^3\right\}.$$

We introduce some notation.

$$\hat{\varphi}(x,y) = \varphi(x') + \varphi(y') - \varphi(x) - \varphi(y),$$

$$Q[\varphi,\psi] = \frac{1}{4}\int_{(0,\pi)\times(0,2\pi)} \hat{\varphi}(x,y)\hat{\psi}(x,y)Q(x,y,\theta)d\theta d\epsilon ,$$

$$\varphi^{\#}(x) = \varphi(x)/e_0(x), \quad \tilde{\varphi}(x,y) = \widehat{\varphi^{\#}}(x,y) .$$

Let $F(\eta)$ be a function of the form

$$F(\eta) = f(\langle\eta,\varphi_1\rangle,\dots,\langle\eta,\varphi_m\rangle), \quad f\in C_0^\infty(R^3), \quad \varphi_1,\dots,\varphi_m\in\mathcal{S},$$

where $\mathcal{S}=\mathcal{S}(R^3)$, the space of rapidly decreasing C^∞-functions. Then the generator L_n of the Markov process $Y_n(t)$ is given by

$$L_nF(\eta) = \sum_{\alpha=1}^{m}\left\langle e_0\eta\otimes\left(\frac{e_0\eta}{\sqrt{n}} + 2g\right) , K^{\#}\varphi_\alpha^{\#}(x,y)\right\rangle\partial_\alpha f$$

$$+ \sum_{\alpha,\beta=1}^{m}\left\langle\left(\frac{e_0\eta}{\sqrt{n}} + g\right)^{2\otimes}, Q[\varphi_\alpha^{\#},\varphi_\beta^{\#}]\right\rangle\partial_{\alpha\beta}^2 f + R_n ,$$

$$R_n = \frac{1}{6n^2\sqrt{n}}\sum_{\alpha,\beta,\gamma=1}^{m}\sum_{1\le i<j\le n}\int_{(0,\pi)\times(0,2\pi)}(\tilde{\varphi}_\alpha\tilde{\varphi}_\beta\tilde{\varphi}_\gamma)(x_i,x_j)\partial_{\alpha\beta\gamma}^3 fQd\theta d\epsilon ,$$

where the argument in $\partial_{\alpha\beta\gamma}^3 f$ is

$$(\langle\eta,\varphi_1\rangle + \frac{\kappa\tilde{\varphi}_1(x_i,x_j)}{\sqrt{n}},\dots,\langle\eta,\varphi_m\rangle + \frac{\kappa\tilde{\varphi}_m(x_i,x_j)}{\sqrt{n}}) , \quad 0<\kappa<1 ,$$

and hence the limiting generator $L=\lim_{n\to\infty}L_n$ is given (at least formally) by

$$LF(\eta) = 2\sum_{\alpha=1}^{m}\langle e_0\eta\otimes g, K^{\#}\varphi_\alpha^{\#}\rangle\partial_\alpha f + \sum_{\alpha,\beta=1}^{m}\langle g\otimes g, Q[\varphi_\alpha^{\#},\varphi_\beta^{\#}]\rangle\partial_{\alpha\beta}^2 f .$$

Now the problem may be stated as follows: (a) Find a diffusion process $Y(t)$ with generator L on a suitable space of distributions. (b) Prove the convergence in the law sense of $Y_n(t)$ to $Y(t)$.

Linearized collision operator: If we introduce l.c.o. $\mathcal{L}$ by

$$\mathcal{L}\varphi(x) = e_0(x)\int_{(0,\pi)\times(0,2\pi)\times R^3} g(y)\tilde{\varphi}(x,y)Q(x,y,\theta)d\theta d\epsilon dy ,$$

L can be expressed as

(9) $$LF(\eta) = - \sum_{\alpha,\beta=1}^{m} (\mathcal{L}\varphi_\alpha, \varphi_\beta)_0 \partial^2_{\alpha\beta} f + \sum_{\alpha=1}^{m} \langle \eta, \mathcal{L}\varphi_\alpha \rangle \partial_\alpha f ,$$

where (,) denotes the usual inner product in $L^2(R^3)$.

<u>Eigenfunctions of $\mathcal{L}$</u>: For each integer $\ell \geq 0$ we choose $2\ell+1$ harmonic polynomials H_ℓ^m, $m=0, \pm 1, \ldots, \pm\ell$ of degree ℓ so that the family

$$\left\{ S_\ell^m \equiv H_\ell^m \big|_{S^2} (\text{spherical harmonics}): \quad m=0, \pm 1, \ldots, \pm\ell \right\}$$

forms an ONS in $L^2(S^2)$, and set

$$e_{\underline{n}}(x) = c_{\underline{n}} e_0(x) L_n^{\ell+\frac{1}{2}}(|x|^2/2) H_\ell^m(x) ,$$

$$\underline{n} = (n,\ell,m), \quad n,\ell = 0,1,\ldots, \quad m = 0,\pm 1,\ldots,\pm\ell ,$$

where

$L_n^{\ell+\frac{1}{2}}(t)$ = the associated Laguerre polynomial of degree n and order $\ell + \frac{1}{2}$,

$$c_{\underline{n}} = \left\{ \frac{n!\pi^{3/2}}{2^{\ell-1}\,\Gamma(n+\ell+\frac{3}{2})} \right\}^{1/2} .$$

Then $\left\{ e_{\underline{n}} : \underline{n} = (n,\ell,m) \right\}$ is a CONS in $L^2(R^3)$. It is known that $e_{\underline{n}}$'s are eigenfunctions of $\mathcal{L}$ ([11] [12]), i.e.,

$$e_{\underline{n}} = -\lambda_{\underline{n}} e_{\underline{n}} ,$$

where

$$\lambda_{\underline{n}} = 2\pi \int_0^\pi \left\{ 1 - (\cos\tfrac{\theta}{2})^{2n+\ell} P_\ell(\cos\tfrac{\theta}{2}) - (\sin\tfrac{\theta}{2})^{2n+\ell} P_\ell(\sin\tfrac{\theta}{2}) \right\} Q(\theta)\, d\theta$$

$$\geq 0 \qquad \text{if} \quad 2n + \ell > 0 ,$$

$$\lambda_{(0,0,0)} = \lambda_{(0,1,m)} = \lambda_{(1,0,0)} = 0 \quad \text{for} \quad m = 0, \pm 1 ,$$

$P_\ell(t)$ being the Legendre polynomial of degree ℓ . We can prove that

(10) $$\lambda_{\underline{n}} \leq \text{const.}(2n + \ell)^2 .$$

Moreover, $e_{\underline{n}}$'s are also eigenfunctions of $\frac{|x|^2}{4} - \Delta$, i.e.,

(11) $$\left(\frac{|x|^2}{4} - \Delta\right)e_{\underline{n}} = \left(2n + \ell + \tfrac{3}{2}\right)e_{\underline{n}} .$$

Spaces of distributions:

$$\|\varphi\|_\alpha = \left\|\left(\frac{|x|^2}{4} - \Delta\right)^{\alpha/2}\varphi\right\|_0 , \quad \alpha \in R$$

$$\left(= \left\{\sum_{\underline{n}} a_{\underline{n}}^2 \left(2n + \ell + \tfrac{3}{2}\right)^\alpha\right\}^{1/2} \text{ for } \varphi = \sum a_{\underline{n}} e_{\underline{n}} \text{ by (11)}\right),$$

$\mathcal{S}_\alpha$ = the Hilbert space obtained by completing $\mathcal{S}$ with respect to $\|\cdot\|_\alpha$,

$\mathcal{S}'_\alpha$ = the dual space of $\mathcal{S}_\alpha$ $(\cong \mathcal{S}_{-\alpha})$.

Stochastic differential equation: Let $B(t)$ be a Brownian motion on $\mathcal{S}'$ determined by

$$E\left\{e^{i\langle B(t),\varphi\rangle}\right\} = e^{-t\|\varphi\|_0^2/2} , \ \varphi \in \mathcal{S} , \quad B(0) = 0;$$

in fact $B(t)$ exists as an $\mathcal{S}'_{3+\varepsilon}$-Brownian motion $(\forall \varepsilon > 0)$. From the expression (9) of $\mathcal{L}$ it follows that the Markov process $Y(t)$ with generator L should satisfy the stochastic differential equation

(12) $$dY(t) = \sqrt{-2\mathcal{L}}\, dB(t) + \mathcal{L}\, Y(t)dt .$$

If we set $Y_{\underline{n}}(t) = \langle Y(t), e_{\underline{n}}\rangle$, then (12) yields

$$dY_{\underline{n}}(t) = \sqrt{2\lambda_{\underline{n}}}\, dB_{\underline{n}}(t) - \lambda_{\underline{n}} Y_{\underline{n}}(t)dt ,$$

where $B_{\underline{n}}(t)$'s are independent copies of 1-dimensional Brownian motion. Making use of the bound (10) for $\lambda_{\underline{n}}$ it can be proved that

$$\sum_{2n+\ell \le k} Y_{\underline{n}}(t)e_{\underline{n}} \to Y(t) \quad \text{as} \quad k \to \infty$$

uniformly on each finite t-interval in the space $\mathcal{S}'_{3+\varepsilon}$ almost surely for any $\varepsilon > 0$, and we can finally prove the following theorem.

Theorem 3. The Markov process $Y(t)$ with generator L and satisfying

$$E\left\{e^{i\langle Y(0),\varphi\rangle}\right\} = e^{-\left\{\|\varphi\|_0^2-(e_0,\varphi)_0^2\right\}/2}, \quad \varphi\in\mathcal{S}$$

exists as a diffusion process on $\mathcal{S}'_{3+\varepsilon}$ for any $\varepsilon>0$, and any finite dimensional distribution of $Y_n(t)$ as a process on $\mathcal{S}'_{3+\varepsilon}$ converges to the corresponding finite dimensional distribution of $Y(t)$ as $n\to\infty$.

It is to be noted that the SDE (12) has a solution $Y(t)$ which is continuous in $\mathcal{S}'_{3+\varepsilon}$, while the process $\sqrt{-2\mathcal{L}}\,B(t)$ is not always continuous in $\mathcal{S}'_{3+\varepsilon}$.

References

1. M. Kac, Foundations of kinetic theory, Proc. Third Berkeley Symp. on Math. Stat. and Prob., 3(1956), 171-197.
2. M. Kac, Some probabilistic aspects of the Boltzmann-equation, Acta Physica Austriaca, Suppl. X(1973), 379-400.
3. H. P. McKean, An exponential formula for solving Boltzmann's equation for a Maxwellian gas, J. Combinatorial Theory, 2(1967), 358-382.
4. H. P. McKean, Fluctuations in the kinetic theory of gases, Comm. Pure Appl. Math., 28(1975), 435-455.
5. F. A. Grünbaum, Propagation of chaos for the Boltzmann equation, Arch. Rational Mech. Anal., 42(1971), 323-345.
6. H. Murata, Propagation of chaos for Boltzmann-like equation of non-cutoff type in the plane, Hiroshima Math. J., 7(1977), 479-515.
7. H. Tanaka, Fluctuation theory for Kac's one-dimensional model of Maxwellian molecules, to appear in Sankhyā, 44, Series A, Part I (1982).
8. K. Uchiyama, Fluctuations of Markovian systems in Kac's caricature of a Maxwellian gas, to appear in J. Math. Soc. Japan.
9. K. Uchiyama, A Fluctuation problem associated with the Boltzmann equation for a gas of molecules with a cut-off potential, to appear.
10. L. Arkeryd, On the Boltzmann equation, Part I, II, Archive for Rational Mech. Anal., 45(1972), 1-16, 17-34.
11. C. S. Wang Chang and G. E. Uhlenbeck (1952), In "The Kinetic Theory of Gases", Studies in Statistical Mechanics, Vol. V, North-Holland, Amsterdam (1970).
12. C. Cercignani, Mathematical Methods in Kinetic Theory, Plenum Press, New York (1969).

Reprinted from

Theory and Application of Random Fields, 258–267, Lecture Notes in Control and Information Sciences, **49**, Springer-Verlag, 1983.

Taniguchi Symp. SA
Katata 1982, pp. 469–488

Limit Theorems for Certain Diffusion Processes with Interaction

Hiroshi TANAKA

Introduction

Given smooth functions $\sigma(x, y)$ and $b(x, y)$, we write

$$\sigma[x, u] = \int_R \sigma(x, y)u(y)dy \quad \left(= \int_R \sigma(x, y)u(dy)\right)$$

$$b[x, u] = \int_R b(x, y)u(y)dy \quad \left(= \int_R b(x, y)u(dy)\right)$$

for a function $u(y)$ (or a measure $u(dy)$). On the analogy of Kac's master equation approach [5] to the Boltzmann equation, McKean [9] considered the n-particle diffusion process $\xi^{(n)}(t) = (\xi_1^{(n)}(t), \cdots, \xi_n^{(n)}(t))$ described by the stochastic differential equation (abbreviated: SDE)

$$(1) \qquad \begin{aligned} d\xi_i^{(n)}(t) = {} & \frac{1}{n-1} \sum_{j=1(\neq i)}^{n} \sigma(\xi_i^{(n)}(t), \xi_j^{(n)}(t))dw_j(t) \\ & + \frac{1}{n-1} \sum_{j=1(\neq i)}^{n} b(\xi_i^{(n)}(t), \xi_j^{(n)}(t))dt, \quad 1 \leq i \leq n, \end{aligned}$$

($w_j(t)$'s are independent copies of a 1-dimensional Brownian motion) and proved the following: If the initial values $\xi_1^{(n)}(0), \cdots, \xi_n^{(n)}(0)$ are independent random variables with the same distribution u, then for each m the process $(\xi_1^{(n)}(t), \cdots, \xi_m^{(n)}(t))$ converges in law to $(\xi_1(t), \cdots, \xi_m(t))$ as $n \to \infty$ where $\xi_1(t), \xi_2(t), \cdots$ are independent copies of the diffusion process $\xi(t)$ obtained by solving the stochastic differential equation

$$(2.a) \qquad d\xi(t) = \sigma[\xi(t), u(t)]dw(t) + b[\xi(t), u(t)]dt$$

(2.b) $\qquad \xi(0)$ is independent of $w(t)$ and u-distributed

subject to the condition "$u(t)$ = the probability distribution of $\xi(t)$", $w(t)$ being a 1-dimensional Brownian motion. Moreover, $u(t)$ is the weak solution of the nonlinear parabolic equation

(3) $$\frac{\partial u}{\partial t} = \frac{1}{2}\cdot\frac{\partial^2}{\partial x^2}(\sigma[x, u]^2 u) - \frac{\partial}{\partial x}(b[x, u]u)$$

with $u(0) = u$. Let W be the space of real valued continuous paths. Then each process $\xi_i^{(n)}(t)$ is regarded as a W-valued random variable $\xi_i^{(n)}$ and McKean's result implies the following law of large numbers:

(4) $$U_n = \frac{1}{n}\sum_{i=1}^{n}\delta_{\xi_i^{(n)}} \longrightarrow \mu, \quad n \to \infty \quad \text{(in prob.)};$$

where μ is the probability measure on W induced by the process $\xi(t)$ and the notation δ_x stands for the δ-distribution at x. The next stage is the central limit theorem in which the asymptotic behavior of

(5) $$Y_n = \sqrt{n}(U_n - \mu),$$

as $n \to \infty$, is studied. The special case when $\sigma = 1$ and $b(x, y) = -\lambda(x - y)$, $\lambda > 0$, was discussed by Tanaka and Hitsuda [14]. We shall discuss, in this paper, the case when $\sigma = 1$ with general $b(x, y)$ but the SDE (1) will be slightly modified as follows:

(6) $$d\xi_i^{(n)}(t) = dw_i(t) + \frac{1}{n}\sum_{j=1}^{n} b(\xi_i^{(n)}(t), \xi_j^{(n)}(t))dt.$$

Braun and Hepp [1] studied similar limit theorems in the case of modified Vlasov equation starting from a (deterministic) n-body problem. Their result on the law of large numbers is covered by McKean's result, since σ may be assumed to vanish in (1). However, the work of Braun and Hepp is very interesting because of their method used in the discussion of the central limit theorem. The purpose of this paper is to study the central limit theorem and the large deviation problem for U_n in the case when $b(x, y)$ is general but $\sigma = 1$, by amplifying the method of Braun and Hepp. The central limit theorem in the time evolution, which studies the asymptotic behavior of Markov processes

(7) $$Y_n(t) = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\delta_{\xi_i^{(n)}(t)} - u(t)\right)$$

as $n \to \infty$, may also be discussed by employing the method of martingale problem as in the works [12] [13] [15] for Boltzmann's equation; however, the result in the time evolution does not automatically imply the result in the path space. The advantage of the present method is that we can easily arrive at the result in the path space and also can find the I-functional governing the large deviation for U_n. This result for the large deviation problem of U_n will lead to a conjecture for a similar problem in the case of Boltzmann's equation.

§ 1. A method of Braun and Hepp

We state a general theorem which is an abstraction of the method used by Braun and Hepp in the proof of Theorem 3.5 of [1].

Let S be a Polish space and $\mathfrak{M}$ be the Banach space of bounded signed measures on S with total variation norm $\|\cdot\|$. Let $\mathfrak{M}_1$ be the subset of $\mathfrak{M}$ consisting of substochastic measures on S. Suppose we are given a function

$$f\colon S \times \mathfrak{M} \longrightarrow \boldsymbol{R}$$

satisfying the following assumption.

Assumption. There exists a (Borel) function

$$f'\colon S \times S \times \mathfrak{M}_1 \longrightarrow \boldsymbol{R}$$

satisfying the following conditions (A.1)–(A.4).

(A.1) f' is bounded.

(A.2) $\|f'(x, y, \rho) - f'(x, y, \rho')\|_\infty \leq \text{const.}\ \|\rho - \rho'\|$ for any $\rho, \rho' \in \mathfrak{M}_1$.

(A.3) If $\rho_n, \rho \in \mathfrak{M}_1$ and if $\rho_n \to \rho$ weakly as $n \to \infty$, then

$$\lim_{n\to\infty} f'(x, y, \rho_n) = f'(x, y, \rho)\,.$$

(A.4) $f(x, \rho+\nu) = f(x, \rho) + f'(x, \nu, \rho) + \int_0^1 \{f'(x, \nu, \rho + t\nu) - f'(x, \nu, \rho)\}dt$

for any $\rho \in \mathfrak{M}_1$ and $\nu \in \mathfrak{M}$ such that $\rho + \nu \in \mathfrak{M}_1$, where

$$f'(x, \nu, \rho) = \int_S f'(x, y, \rho)\nu(dy)\,.$$

Notation. For $\rho, \rho' \in \mathfrak{M}$ we use the notation

$$f(\rho', \rho) = \int_S f(x, \rho)\rho'(dx)\,.$$

Theorem 1.1. *Let f be a bounded function satisfying the above assumption and $X_1, X_2, \cdots$ be a sequence of independent S-valued random variables with the same distribution λ. We set*

$$\lambda_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}\,,$$

$$Y_n = \sqrt{n}\{f(\lambda_n, \lambda_n) - f(\lambda, \lambda)\} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{f(X_i, \lambda_n) - f(\lambda, \lambda)\}.$$

Then the probability distribution of Y_n converges to the Gaussian distribution with mean 0 *and variance σ^2 as $n \to \infty$, where*

$$\sigma^2 = \int_S \{f(x, \lambda) + f'(\lambda, x, \lambda) - m\}^2 \lambda(dx),$$

$$m = \int_S \{f(x, \lambda) + f'(\lambda, x, \lambda)\}\lambda(dx).$$

Proof. First we consider the *special case* when $f(x, \rho)$ is given by

$$f(x, \rho) = \int_S f(x, y)\rho(dy)$$

with a bounded function $f(x, y)$ on $S \times S$ (for simplicity we use the same notation with confusion). In this case $f'(x, y, \rho) = f(x, y)$ and hence the variance σ^2 is equal to

$$\int_S \{f(x, \lambda) + f(\lambda, x) - 2f(\lambda, \lambda)\}^2 \lambda(dx).$$

If we set

$$\begin{aligned}\tilde{Y}_n &= \sqrt{n}\{f(\lambda_n, \lambda) + f(\lambda, \lambda_n) - 2f(\lambda, \lambda)\} \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{f(X_i, \lambda) + f(\lambda, X_i) - 2f(\lambda, \lambda)\},\end{aligned}$$

then the probability distribution of $\tilde{Y}_n$ converges to $N(0, \sigma^2)$ as $n \to \infty$, and hence it is enough to prove that

$$E\{|Y_n - \tilde{Y}_n|^2\} \longrightarrow 0, \quad n \to \infty.$$

Setting $Z_{ni} = f(X_i, \lambda_n) - f(X_i, \lambda) - f(\lambda, \lambda_n) + f(\lambda, \lambda)$, we have

$$\begin{aligned}Y_n - \tilde{Y}_n &= \sqrt{n}\{f(\lambda_n, \lambda_n) - f(\lambda, \lambda_n) - f(\lambda, \lambda_n) + f(\lambda, \lambda)\} \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_{ni}.\end{aligned}$$

We now claim

(i) $E\{Z_{ni}^2\} = o(1), \quad n \to \infty,$

(ii) $E\{Z_{ni}Z_{nj}\} = O(n^{-2}), \quad n \to \infty \quad (i \neq j).$

(i) is immediate from the law of large numbers. As for (ii), setting

$$\lambda_n^{ij} = \frac{1}{n} \sum_{k=1(\neq i,j)}^{n} \delta_{X_k},$$

$$Z_{nij} = f(X_i, \lambda_n^{ij}) - f(X_i, \lambda) - f(\lambda, \lambda_n^{ij}) + f(\lambda, \lambda),$$
$$R_{nij} = f(X_i, X_i) + f(X_i, X_j) - f(\lambda, X_i) - f(\lambda, X_j),$$

we have $Z_{ni} = Z_{nij} + n^{-1}R_{nij}$ $(i \neq j)$, and hence

$$E\{Z_{ni}Z_{nj}\} = E\{Z_{nij}Z_{nji}\} + n^{-1}E\{Z_{nij}R_{nji}\} + n^{-1}E\{Z_{nji}R_{nij}\} + 0(n^{-2}).$$

The first term on the right hand side of the above vanishes because

$$E\{Z_{nij}Z_{nji} \mid X_k, k \neq i, j\} = 0,$$

while the second and the third terms are $O(n^{-2})$ because

$$\begin{aligned} E\{Z_{nij}R_{nji} \mid X_i, X_j\} &= E[R_{nji}E\{Z_{nij} \mid X_i, X_j\}] \\ &= E[R_{nji}\{n^{-1}(n-2)f(X_i, \lambda) - f(x_i, \lambda) \\ &\qquad - n^{-1}(n-2)f(\lambda, \lambda) + f(\lambda, \lambda)\}] \\ &= E[R_{nji} \cdot 2n^{-1}\{f(\lambda, \lambda) - f(X_i, \lambda)\}] = O(n^{-1}). \end{aligned}$$

Since $E\{Z_{ni}^2\}$ and $E\{Z_{ni}Z_{nj}\}$ are independent of i and j $(i \neq j)$, (i) and (ii) imply

$$\lim_{n\to\infty} E\{|Y_n - \tilde{Y}_n|^2\} = \lim_{n\to\infty} n^{-1} \sum_{i,j=1}^{n} E\{Z_{ni}Z_{nj}\} = 0.$$

Next we consider the *general case*. If we set

$$\tilde{Y}_n = \sqrt{n}\{f(\lambda_n, \lambda) + f(\lambda, \lambda_n) - 2f(\lambda, \lambda)\}$$
$$Z_{ni} = f(X_i, \lambda_n) - f(X_i, \lambda) - f(\lambda, \lambda_n) + f(\lambda, \lambda),$$

then

$$\begin{aligned} Y_n - \tilde{Y}_n &= \sqrt{n}\{f(\lambda_n, \lambda_n) - f(\lambda_n, \lambda) - f(\lambda, \lambda_n) + f(\lambda, \lambda)\} \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_{ni}. \end{aligned}$$

Step 1. $E\{|Y_n - \tilde{Y}_n|^2\} \to 0$, $n \to \infty$. For this it is enough to prove the followings:

(iii) $E\{Z_{ni}^2\} = o(1), \quad n \to \infty,$

(iv) $E\{Z_{ni}Z_{nj}\} = O(n^{-2}), \quad n \to \infty \quad (i \neq j).$

Setting $\rho = \lambda$ and $\nu = \lambda_n - \lambda$ in (A.4), we have

$$f(x, \lambda_n) = f(x, \lambda) + f'(x, \lambda_n - \lambda, \lambda) + \int_0^1 \{f'(x, \lambda_n - \lambda, \lambda + t(\lambda_n - \lambda)) - f'(x, \lambda_n - \lambda, \lambda)\}dt ,$$

and hence

$$f(x, \lambda_n) - f(x, \lambda) = \varphi(x, \lambda_n - \lambda) + \varphi_n(x, \lambda_n - \lambda, \omega) = n^{-1} \sum_{j=1}^{n} \varphi(x, X_j) + n^{-1} \sum_{j=1}^{n} \varphi_n(x, X_j, \omega) ,$$

where

$$\varphi(x, y) = f'(x, y, \lambda) - f'(x, \lambda, \lambda)$$

$$\varphi_n(x, y, \omega) = \int_0^1 \{f'(x, y, \lambda + t(\lambda_n - \lambda)) - f'(x, y, \lambda)\}dt - \int_0^1 \{f'(x, \lambda, \lambda + t(\lambda_n - \lambda)) - f'(x, \lambda, \lambda)\}dt .$$

Therefore

$$\begin{aligned} E\{|f(X_i, \lambda_n) - f(X_i, \lambda)|^2\} &\leq 2n^{-2} \sum_{j,k=1}^{n} E\{\varphi(X_i, X_j)\varphi(X_i, X_k)\} \\ &\quad + 2n^{-2} \sum_{j,k=1}^{n} E\{\varphi_n(X_i, X_j, \omega)\varphi_n(X_i, X_k, \omega)\} \\ &= 2n^{-2} \sum{}' E\{\varphi(X_i, X_j)\varphi(X_i, X_k)\} \\ &\quad + 2n^{-2} \sum{}' E\{\varphi_n(X_i, X_j, \omega)\varphi_n(X_i, X_k, \omega)\} + O(n^{-1}) , \end{aligned}$$

where $\sum'$ is the summation over all pairs (j, k) such that $1 \leq j, k \leq n$ and $j \neq i$, $k \neq i$, $j \neq k$. Since the first term in the above vanishes and since the number of such pairs (j, k) is $(n - 1)(n - 2)$, we have

$$\begin{aligned} E\{|f(X_i, \lambda_n) - f(X_i, \lambda)|^2\} &\leq 2n^{-2}(n - 1)(n - 2)E\{\varphi_n(X_1, X_2, \omega)\varphi_n(X_1, X_3, \omega)\} + O(n^{-1}) \\ &\longrightarrow 0 , \quad n \to \infty , \end{aligned}$$

because $\varphi_n(x, y, \omega) \to 0$ as $n \to \infty$ (a.s.) for each fixed $x, y \in S$ by (A.3). Similarly

$$E\{|f(\lambda, \lambda_n) - f(\lambda, \lambda)|^2\} \longrightarrow 0 , \quad n \to \infty ,$$

and hence we obtain (iii). For the proof of (iv) let $i \neq j$. Noticing that

$$f(X_i, \lambda_n) = f(X_i, \lambda_n^{ij}) + n^{-1}\{f'(X_i, X_i, \lambda_n^{ij}) + f'(X_i, X_j, \lambda_n^{ij})\} + O(n^{-2})$$

and also a similar formula for $f(\lambda, \lambda_n)$, we have

$$Z_{ni} = Z_{nij} + n^{-1}R_{nij}\,,$$

where

$$\begin{aligned} Z_{nij} &= f(X_i, \lambda_n^{ij}) - f(X_i, \lambda) - f(\lambda, \lambda_n^{ij}) + f(\lambda, \lambda)\,, \\ R_{nij} &= f'(X_i, X_i, \lambda_n^{ij}) + f'(X_i, X_j, \lambda_n^{ij}) \\ &\qquad - f'(\lambda, X_i, \lambda_n^{ij}) - f'(\lambda, X_j\, \lambda_n^{ij})\,. \end{aligned}$$

Therefore

$$E\{Z_{ni}Z_{nj}\} = E\{Z_{nij}Z_{nji}\} + n^{-1}E\{Z_{nij}R_{nji} + Z_{nji}R_{nij}\} + O(n^{-2})$$

which is still $O(n^{-2})$ as $n \to \infty$, because the first and the second terms on the right hand side of the above vanish as will be verified below:

$$\begin{aligned} E\{Z_{nij}Z_{nji}\} &= E[E\{Z_{nij}Z_{nji} \mid X_k, k \neq i, j\}] = 0\,, \\ E\{Z_{nij}R_{nji}\} &= E[Z_{nij}\{f'(X_j, X_i, \lambda_n^{ij}) - f'(\lambda, X_i, \lambda_n^{ij})\}] \\ &\quad + E[Z_{nij}\{f'(X_j, X_j, \lambda_n^{ij}) - f'(\lambda, X_j, \lambda_n^{ij})\}] \\ &= E[Z_{nij}E\{f'(X_j, X_i, \lambda_n^{ij}) - f'(\lambda, X_i, \lambda_n^{ij}) \mid X_k, k \neq j\}] \\ &\quad + E[\{f'(X_j, X_j, \lambda_n^{ij}) - f'(\lambda, X_j, \lambda_n^{ij})\}E\{Z_{nij} \mid X_k, k \neq i\}] \\ &= 0\,. \end{aligned}$$

Step 2. *The distribution of $\tilde{Y}_n$ converges to $N(0, \sigma^2)$.* Using the notation $\varphi_n(x, y, \omega)$ in the proof of Step 1, we can write

$$f(\lambda, \lambda_n) - f(\lambda, \lambda) = f'(\lambda, \lambda_n - \lambda, \lambda) + \varphi_n(\lambda, \lambda_n, \omega)\,,$$

and hence

$$\begin{aligned} \tilde{Y}_n &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{f(X_i, \lambda) - f(\lambda, \lambda) + f(\lambda, \lambda_n) - f(\lambda, \lambda)\} \\ &= \sqrt{n}\,\varphi(\lambda_n - \lambda) + \sqrt{n}\,\varphi_n(\lambda, \lambda_n, \omega)\,, \end{aligned}$$

where $\varphi(x) = f(x, \lambda) + f'(\lambda, x, \lambda)$. Since the distribution of $\sqrt{n}\,\varphi(\lambda_n - \lambda)$ converges to $N(0, \sigma^2)$, it is enough to prove that

$$E\{|\sqrt{n}\,\varphi_n(\lambda, \lambda_n, \omega)|^2\} = o(1)\,, \quad n \to \infty\,.$$

The left hand side of the above equals

$$E\left\{\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\varphi_n(\lambda, X_i, \omega)\right|^2\right\}$$
$$= E\left\{\left|\varphi_n(\lambda, X_1, \omega)\right|^2\right\} + \frac{n^2 - n}{n} E\{\varphi_n(\lambda, X_1, \omega)\varphi_n(\lambda, X_2, \omega)\} .$$

The first term of the above tends to 0 as $n \to \infty$, because $\varphi_n(\lambda, X_1, \omega) \to 0$ boundedly (a.s.) by (A.3). As for the second term we first write

$$\varphi_n(\lambda, y, \omega) = \psi_n(\lambda, y, \omega) + O(n^{-1}) \quad \text{(by (A.2))} ,$$
$$\psi_n(\lambda, y, \omega) = \int_0^1 \{f'(\lambda, y, \lambda + t(\lambda_n^{12} - \lambda)) - f'(\lambda, y, \lambda)\}dt$$
$$+ \int_0^1 \{f'(\lambda, \lambda, \lambda + t(\lambda_n^{12} - \lambda)) - f'(\lambda, \lambda, \lambda)\}dt ,$$

and then notice that

$$E\{\psi_n(\lambda, X_1, \omega)\psi_n(\lambda, X_2, \omega) \,|\, X_k,\ k \neq 1, 2\} = 0 .$$

Thus the second term also tends to 0 as $n \to \infty$, because

$$E\{\varphi_n(\lambda, X_1, \omega)\varphi_n(\lambda, X_2, \omega)\}$$
$$= E\{\psi_n(\lambda, X_1, \omega)\psi_n(\lambda, X_2, \omega)\}$$
$$+ \frac{\text{const.}}{n} E\{|\psi_n(\lambda, X_1, \omega)| + |\psi_n(\lambda, X_2, \omega)|\} + O(n^{-2})$$
$$= o(n^{-1}) \quad \text{(by (A.3))} .$$

§ 2. Central limit theorem

In this section we explain how Theorem 1.1 can be applied to the central limit theorem for Y_n of (5) when the coefficient $\sigma = 1$.

2.1. Let W be the space of continuous paths $w: t \in [0, 1] \to R$. For simplicity the time parameter t is restricted to $0 \leq t \leq 1$. W is a Banach space with the maximum norm $\|\cdot\|_\infty$ and will stands for the Polish space S of § 1. Let $\mathfrak{M}$ be the Banach space of finite signed measures on W with total variation norm $\|\cdot\|$, $\mathfrak{M}_1$ be the subset of $\mathfrak{M}$ consisting of substochastic measures on W and $C_b^2(\boldsymbol{R}^2)$ be the space of C^2-functions which are bounded together with their first and second partial derivatives. Given a coefficient $b = b(x, y) \in C_b^2(\boldsymbol{R}^2)$ and a signed measure $\rho \in \mathfrak{M}$, we consider the equation

$$\text{(2.1)} \qquad \xi(t, w) = w(t) + \int_0^t b_\rho(\xi(s, w), \xi(s))ds , \quad 0 \leq t \leq 1 ,$$

where $\xi(s)$ is the map: $w \in W \to \xi(s, w) \in \boldsymbol{R}$ and the notation

$$b_\rho(x, \alpha) = \int_W b(x, \alpha(w))\rho(dw)$$

is used for a map $\alpha\colon W \to \boldsymbol{R}$. The equation (2.1) can be solved easily by iteration. We denote by $\xi(t, w, \rho)$ the solution of (2.1). When we regard $\xi(t, w, \rho)$ as a function of t alone by fixing w and ρ, we denote it by $\xi(w, \rho)$. Then $\xi(w, \rho) \in W$, because

$$(2.2) \qquad |\xi(t, w, \rho) - \xi(s, w, \rho)| \le |w(t) - w(s)| + \|\rho\| \cdot \|b\|_\infty \cdot |t - s|\,.$$

Lemma 2.1. *$\xi(w, \rho)$ has the following properties.*

(i) $\|\xi(w, \rho) - \xi(w', \rho)\|_\infty \le c_1 \|w - w'\|_\infty$, *where c_1 is a constant depending only on $\|\rho\|$ and $\|b_x\|_\infty$ ($b_x = \partial b/\partial x$, etc.).*

(ii) $\|\xi(w, \rho) - \xi(w, \rho')\|_\infty \le c_2 \|\rho - \rho'\|$, *where c_2 is a constant depending only on $\|\rho\|$, $\|\rho'\|$, $\|b\|_\infty$, $\|b_x\|_\infty$ and $\|b_y\|_\infty$.*

(iii) *Let $\rho_n, \rho \in \mathfrak{M}_1$ and $\rho_n \to \rho$ weakly as $n \to \infty$. Then*

$$\xi(w, \rho_n) \to \xi(w, \rho) \quad (in\ W), \quad n \to \infty$$

uniformly on each compact subset K of W.

Proof. We give only the proof of (iii). By (2.2) and (i) the family of functions $\xi(t, w, \rho_n)$, $n = 1, 2, \cdots$, on $[0, 1] \times K$ is uniformly bounded and equicontinuous provided that K is a compact subset of W. Therefore, for any subsequence $n_1 < n_2 < \cdots$ we can choose a further subsequence $\{n'_k\}$ of $\{n_k\}$ such that

(2.3a) $\xi(t, w, \rho_{n'_k})$ converges to some $\xi(t, w)$ as $k \to \infty$ for any (t, w),

(2.3b) the above convergence is uniform on $[0, 1] \times K$ for each compact subset K of W.

The assertion (iii) follows once we identify $\xi(t, w)$ as $\xi(t, w, \rho)$. This can be done as follows. Letting $n \to \infty$ via $\{n'_k\}$ in

$$\xi(t, w, \rho_n) = w(t) + \int_0^t ds \int_W \rho_n(dw') b(\xi(s, w, \rho_n), \xi(s, w', \rho_n))\,,$$

we have

$$\xi(t, w) = w(t) + \int_0^t ds \int_W \rho(dw') b(\xi(s, w), \xi(s, w'))\,,$$

and hence the uniqueness of the solution implies $\xi(t, w) = \xi(t, w, \rho)$.

For $\rho, \nu \in \mathfrak{M}$ we set

$$\eta(t, w, \nu, \rho) = \lim_{h\to 0} \frac{1}{h}\{\xi(t, w, \rho + h\nu) - \xi(t, w, \rho)\}.$$

Under the assumption $b \in C_b^2(\boldsymbol{R}^2)$ it can be proved that the above convergence is uniform in $(t, w) \in [0, 1] \times W$, and hence

$$\eta(w, \nu, \rho) = \lim_{h\to 0} \frac{1}{h}\{\xi(w, \rho + h\nu) - \xi(w, \rho)\}$$

also exists as the strong limit in the Banach space W. Moreover, for fixed ρ and ν, $\eta(t) \equiv \eta(t, w) \equiv \eta(t, w, \nu, \rho)$ can be regarded as a W-valued function of t and satisfies

$$\begin{aligned}\dot{\eta}(t, w) = \int_W \rho(dw')\{&b_x(\xi(t, w, \rho), \xi(t, w', \rho))\eta(t, w) \\ &+ b_y(\xi(t, w, \rho), \xi(t, w', \rho))\eta(t, w')\} \\ + \int_W &\nu(dw')b(\xi(t, w, \rho), \xi(t, w', \rho))\end{aligned}$$

with initial condition $\eta(0) = 0$ (the dot $= d/dt$). Setting

$$\eta(t, w, \tilde{w}, \rho) = \eta(t, w, \delta_{\tilde{w}}, \rho), \qquad \eta(w, \tilde{w}, \rho) = \eta(w, \delta_{\tilde{w}}, \rho),$$

we have

$$\begin{aligned}\dot{\eta}(t, w, \tilde{w}, \rho) = \int_W \rho(dw')\{&b_x(\xi(t, w, \rho), \xi(t, w', \rho))\eta(t, w, \tilde{w}, \rho) \\ &+ b_y(\xi(t, w, \rho), \xi(t, w', \rho))\eta(t, w', \tilde{w}, \rho)\} \\ + b(\xi(t, &w, \rho), \xi(t, \tilde{w}, \rho)),\end{aligned} \tag{2.4}$$

$$\eta(t, w, \nu, \rho) = \int_W \nu(d\tilde{w})\eta(t, w, \tilde{w}, \rho). \tag{2.5}$$

Now looking carefully at (2.4) which is a differential equation in the Banach space W (for fixed $\tilde{w}$ and ρ), we see that $\eta(t, w, \tilde{w}, \rho)$ can be expressed as

$$\eta(t, w, \tilde{w}, \rho) = \zeta(t, \xi(w, \rho), \xi(\tilde{w}, \rho), \rho\circ\xi(\rho)^{-1}), \tag{2.6}$$

where $\zeta(t, w, \tilde{w}, \rho)$ is the solution of

$$\begin{aligned}\dot{\zeta}(t, w, \tilde{w}, \rho) = \int_W &\rho(dw')\{b_x(w(t), w'(t))\zeta(t, w, \tilde{w}, \rho) \\ &+ b_y(w(t), w'(t))\zeta(t, w', \tilde{w}, \rho)\} + b(w(t), \tilde{w}(t))\end{aligned} \tag{2.7}$$

with the initial condition $\zeta(0, w, \tilde{w}, \rho) = 0$ and $\rho \circ \xi(\rho)^{-1}$ is the image measure of ρ under the map $\xi(\rho)\colon w \in W \to \xi(w, \rho) \in W$.

When we regard $\zeta(t, w, \tilde{w}, \rho)$ as a function of t alone by fixing w, $\tilde{w}$ and ρ, we denote it by $\zeta(w, \tilde{w}, \rho)$. The proof of the following lemma is easy (especially, the proof of (viii) is similar to that of (iii)), so is omitted.

Lemma 2.2. *$\zeta(w, \tilde{w}, \rho)$ has the following properties.*

(iv) $\|\zeta(w, \tilde{w}, \rho)\|_\infty \le c_3$ *with a constant c_3 depending only on* $\|\rho\|$, $\|b\|_\infty$, $\|b_x\|_\infty$ *and* $\|b_y\|_\infty$.

(v) $\|\zeta(w, \tilde{w}, \rho) - \zeta(w', \tilde{w}, \rho)\|_\infty \le c_4\|w - w'\|_\infty$ *with a constant c_4 depending only on $\|\rho\|$ and the supremum norms of b and its first and second partial derivatives.*

(vi) $\|\zeta(w, \tilde{w}, \rho) - \zeta(w, \hat{w}, \rho)\|_\infty \le c_5\|\tilde{w} - \hat{w}\|_\infty$ *with a constant c_5 depending only on* $\|\rho\|$, $\|b_x\|_\infty$ *and* $\|b_y\|_\infty$.

(vii) $\|\zeta(w, \tilde{w}, \rho) - \zeta(w, \tilde{w}, \rho')\|_\infty \le c_6\|\rho - \rho'\|$ *with a constant c_6 depending only on* $\|\rho\|$, $\|\rho'\|$, $\|b\|_\infty$, $\|b_x\|_\infty$ *and* $\|b_y\|_\infty$.

(viii) *Let $\rho_n, \rho \in \mathfrak{M}_1$ and $\rho_n \to \rho$ weakly as $n \to \infty$. Then*

$$\zeta(w, \tilde{w}, \rho_n) \longrightarrow \zeta(w, \tilde{w}, \rho) \quad (in\ W), \quad n \to \infty$$

uniformly on each compact subset of $W \times W$.

Lemma 2.3. *$\eta(w, \tilde{w}, \rho)$ has the following properties.*

(ix) $\|\eta(w, \tilde{w}, \rho) - \eta(w, \tilde{w}, \rho')\|_\infty \le c_7\|\rho - \rho'\|$ *with a constant c_7 depending only on $\|\rho\|$, $\|\rho'\|$ and the supremum norms of b and its first and second partial derviatives.*

(x) *Let $\rho_n, \rho \in \mathfrak{M}$ and $\rho_n \to \rho$ weakly as $n \to \infty$. Then*

$$\eta(w, \tilde{w}, \rho_n) \to \eta(w, \tilde{w}, \rho_n) \quad (in\ W), \quad n \to \infty$$

uniformly on each compact subset of $W \times W$.

Proof. (ix) can be proved directly from (2.4), and (x) follows from (2.6), (iii) and (viii).

2.2. Given a probability measure u in $\boldsymbol{R}$, we denote by λ the Wiener measure on W such that $\lambda\{w(0) \in \cdot\} = u(\cdot)$. Then the solution $\xi(t)$ of (2) with $\sigma = 1$ is expressed as

$$\xi(t) = \xi(t, w, \lambda). \tag{2.8}$$

To describe the solution of (6) let $\Omega = W \times W \times \cdots$ and P be the product measure $\lambda \otimes \lambda \otimes \cdots$. Then, the solution process $\xi^{(n)}(t) = (\xi_1^{(n)}(t), \cdots, \xi_n^{(n)}(t))$ of (6) with the initial condition

(2.9) $\xi_1^{(n)}(0), \cdots, \xi_n^{(n)}(0)$ are independent random variables with the same distribution u

is realized on the probability space (Ω, P) by

$$\xi_i^{(n)}(t, \omega) = \xi(t, w_i, \lambda_n), \quad 1 \leq i \leq n, \tag{2.10}$$

where $\lambda_n = n^{-1} \sum_{k=1}^{n} \delta_{w_k}$ and $\omega = (w_1, w_2, \cdots)$. Therefore, Y_n of (5) is expressed as

$$Y_n = \sqrt{n}\left\{\frac{1}{n}\sum_{k=1}^{n} \delta_{\xi(w_k, \lambda_n)} - \mu\right\}, \quad \mu = \lambda \circ \xi(\lambda)^{-1}, \tag{2.11}$$

and hence for a (nice) function φ on W we have

$$Y_n(\varphi) = \sqrt{n}\left\{\frac{1}{n}\sum_{k=1}^{n} \varphi(\xi(w_k, \lambda_n)) - \langle \mu, \varphi \rangle\right\}, \tag{2.12}$$

to which Theorem 1.1 can now be applied with $f(w, \rho) = \varphi(\xi(w, \rho))$. A real valued function φ on W is said to be in the class $C_b^1(W)$ if there exists a bounded continuous map φ': $W \to W^*$ (the dual space) such that

$$|\varphi(w + w') - \varphi(w) - \varphi'(w)(w')| = o(\|w'\|_\infty), \quad w' \to 0.$$

For $\varphi \in C_b^1(W)$ we set

$$\Lambda\varphi(\tilde{w}) = E_\mu\{\varphi'(w)(\zeta(w, \tilde{w}, \mu))\}, \quad \tilde{w} \in W, \tag{2.13}$$

$$Q(\varphi) = E_\mu\{|(I + \Lambda)\varphi|^2\} - |E_\mu\{(I + \Lambda)\varphi\}|^2, \tag{2.14}$$

where E_μ denotes the expectation with respect to $\mu(dw)$. If $f(w, \rho) = \varphi(\xi(w, \rho))$, then $f(w, \rho)$ satisfies (A.1)–(A.4) of § 1 with $f'(w, \tilde{w}, \lambda) = \varphi'(\xi(w, \lambda))(\zeta(\xi(w, \lambda), \xi(\tilde{w}, \lambda), \mu))$ and hence by Theorem 1.1 we have the following theorem.

Theorem 2.1. *Let $\xi^{(n)}(t) = (\xi_1^{(n)}(t), \cdots, \xi_n^{(n)}(t))$ be the solution of* (6) *with the initial condition* (2.9) *and let Y_n be defined by* (5) *in the introduction. Then*

$$\lim_{n \to \infty} E\{e^{iY_n(\varphi)}\} = e^{-Q(\varphi)/2}, \quad \varphi \in C_b^1(W), \tag{2.15}$$

that is, the distribution of $Y_n(\varphi)$ converges to $N(0, Q(\varphi))$.

§ 3. Another expression of Λ

The soluution $\xi(t)$ of (2) ($\sigma = 1$) is a temporally inhomogeneous Markov process. This Markov process is called the diffusion process associated with the nonlinear parabolic equation (3) ($\sigma = 1$) ([8]). Similarly, for each fixed $w \in W$ we can consider the diffusion process $\tilde{\xi}(t)$ associated with the nonlinear parabolic equation

$$\frac{\partial \tilde{u}(t)}{\partial t} = \frac{1}{2} \cdot \frac{\partial^2 \tilde{u}(t)}{\partial x^2} - \frac{\partial}{\partial x}[\{b[x, \tilde{u}(t)] + \varepsilon b(x, w(t))\}\tilde{u}(t)] . \tag{3.1}$$

This diffsion process $\tilde{\xi}(t)$ with initial distribution u is described by

$$\tilde{\xi}(t, w') = w'(t) + \int_0^t b_\lambda(\tilde{\xi}(s, w'), \tilde{\xi}(s))ds + \varepsilon \int_0^t b(\tilde{\xi}(s, w'), w(s))ds , \tag{3.2}$$

where $\{w'(t), 0 \leq t \leq 1, \lambda\}$ is a Brownian motion with initial distribution u as before. Noting that $\tilde{\xi}(t) = \xi(t)$ when $\varepsilon = 0$, we can prove by a routine argument in differential equations that

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\{\tilde{\xi}(t, w') - \xi(t, w')\} = \tilde{\eta}(t, w') \tag{3.3}$$

exists as a uniform limit on $[0, 1] \times W$ and that $\tilde{\eta}(t, w')$ satisfies

$$\begin{aligned} \tilde{\eta}(t, w') = \int_0^t ds \int_W \lambda(d\tilde{w})\{& b_x(\xi(s, w')\xi(s, \tilde{w}))\tilde{\eta}(s, w') \\ &+ b_y(\xi(s, w'), \xi(s, \tilde{w}))\tilde{\eta}(s, \tilde{w})\} \\ + \int_0^t & b(\xi(s, w'), w(s))ds . \end{aligned} \tag{3.4}$$

Comparing (3.4) with (2.7) we see that

$$\tilde{\eta}(t, w') = \zeta(t, \xi(w'), w, \mu) , \quad \mu = \lambda \circ \xi(\lambda)^{-1} . \tag{3.5}$$

Theorem 3.1. *For each fixed* $w \in W$ *let* $\mu_{w,\varepsilon}$ *be the probability measure on* W *induced by the diffusion process associated with* (3.1) *having initial distribution* u *and denote by* $E_{\mu,w,\varepsilon}$ *the expectation with respect to* $\mu_{w,\varepsilon}$. *Then for* $\varphi \in C_b^1(W)$

$$\Lambda\varphi(w) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\{E_{\mu,w,\varepsilon}(\varphi) - E_\mu(\varphi)\} . \tag{3.6}$$

Proof. From (2.13), (3.5) and (3.3) we have

$$\begin{aligned} \Lambda\varphi(w) &= \int_W \mu(dw')\varphi'(w')(\zeta(w', w, \mu)) \\ &= \int_W \lambda(dw')\varphi'(\xi(w'))(\tilde{\eta}(w')) \end{aligned}$$

$$= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_W \lambda(dw')\{\varphi(\tilde{\xi}(w')) - \varphi(\xi(w'))\}$$

$$= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \{E_{\mu,w,\varepsilon}(\varphi) - E_\mu(\varphi)\}.$$

§ 4. An infinite dimensional SDE

Let $\{Y(\varphi),\ \varphi \in C_b^1(W)\}$ be a Gaussian system defined on a probability space $(\Omega, \mathscr{F}, P)$ such that

$$E\{e^{iY(\varphi)}\} = e^{-Q(\varphi)/2}, \tag{4.1}$$

$$Y(c_1\varphi_1 + c_2\varphi_2) = c_1 Y(\varphi_1) + c_2 Y(\varphi_2) \quad \text{a.s.}, \tag{4.2}$$

and define $Y(t)$ by $\langle Y(t), f\rangle = Y(f(w(t)))$ for f in the space $\mathscr{S}$ of rapidly decreasing C^∞-functions on $\boldsymbol{R}$. We also define $Y_n(t)$ by (7) or equivalently by $\langle Y_n(t), f\rangle = Y_n(f(w(t))), f \in \mathscr{S}$, where Y_n is given by (5). Then Theorem 2.1 implies the following *central limit theorem in the time evolution scheme*: For any $0 \le t_1 < \cdots < t_m$ and $f_1, \cdots, f_m \in \mathscr{S}$ the joint distribution of $\langle Y_n(t_1), f_1\rangle, \cdots, \langle Y_n(t_m), f_m\rangle$ converges to that of $\langle Y(t_1), f_1\rangle, \cdots, \langle Y(t_m), f_m\rangle$ as $n \to \infty$.

SDE for $Y(t)$: Once we know (2.15), we can derive the SDE for $Y(t)$ as in Itô [4]. Let $u(t)$ be the (probability measure) solution of (3) with $\sigma = 1$ and with initial distribution u. We set

$$(K_{u(t)}f)(x) = \frac{1}{2} f''(x) + b[x, u(t)]f'(x),$$

$$(L_{u(t)}f)(x) = \int b(y, x) f'(y) u(t, dy),$$

$$M_f(t) = f(w(t)) - f(w(0)) - \int_0^t (K_{u(s)}f)(w(s))ds,$$

$$N_f(t) = M_f(t) - \int_0^t (L_{u(s)}f)(w(s))ds.$$

Then $M_f(t)$ is a μ-martingale and it is not hard to prove that

$$E_\mu\{|M_f(t) - M_f(s)|^2 | \mathscr{B}_s\} = \int_s^t \|f'\|_{u(\tau)}^2 d\tau, \quad \text{a.s.} \quad (0 \le s \le t), \tag{4.3}$$

where $\mathscr{B}_s = \sigma\{w(\tau): 0 \le \tau \le s\}$ and $\|\cdot\|_{u(\tau)}$ is the norm in $L^2(R, u(\tau))$.

Lemma 4.1. $(I + \Lambda)N_f(t) = M_f(t), \quad f \in \mathscr{S}$.

Proof. It is enough to prove that

$$
\begin{aligned}
&-\int_0^t (L_{u(s)}f)(w(s))ds + E_\mu\{f'(w(t))\zeta(t, w)\} \\
(4.4) \qquad &-\int_0^t E_\mu\{(K_{u(s)}f)'(w(s))\zeta(s, w)\}ds \\
&-\int_0^t E_\mu\{(L_{u(s)}f)'(w(s))\zeta(s, w)\}ds = 0
\end{aligned}
$$

for each fixed $\tilde{w} \in W$, where $\zeta(t, w) = \zeta(t, w, \tilde{w}, \mu)$. Since $w(t)$ (as a process defined on the probability space (W, μ)) can be expressed as

$$\textit{a Brownian motion} + \int_0^t b[w(s), u(s)]ds\,,$$

an application of Itô's formula yields

$$E_\mu\{f'(w(t))\,\zeta(t, w)\} = E_\mu\left\{\int_0^t F(s)ds\right\},$$

where $F(s)$ is equal to

$$
\begin{aligned}
&f''(w(s))\zeta(s, w)b[w(s), u(s)] + f'(w(s))\dot{\zeta}(s, w) + \frac{1}{2}f'''(w(s))\zeta(s, w) \\
&\quad = b[w(s), u(s)]f''(w(s))\zeta(s, w) + \frac{1}{2}f'''(w(s))\zeta(s, w) \\
&\qquad + f'(w(s))\Big[\int \mu(dw')\{b_x(w(s), w'(s))\zeta(s, w) \\
&\qquad + b_y(w(s), w'(s))\zeta(s, w')\} + b(w(s), \tilde{w}(s))\Big] \quad (by\ (2.7))\,.
\end{aligned}
$$

Therefore, writing down $(K_{u(s)}f)'$ and $(L_{u(s)}f)'$ explicitly, we see that the left hand side of (4.4) is equal to $\int_0^t E_\mu\{G(s)\}ds$ where

$$
\begin{aligned}
G(s) = &-b(w(s), \tilde{w}(s))f'(w(s)) + b[w(s), u(s)]f''(w(s))\zeta(s, w) \\
&+ \frac{1}{2}f'''(w(s))\zeta(s, w) + f'(w(s))\Big[\int \mu(dw')\{b_x(w(s), w'(s))\zeta(s, w) \\
&\qquad + b_y(w(s), w'(s))\zeta(s, w')\} + b(w(s), \tilde{w}(s))\Big] \\
&- \frac{1}{2}f'''(w(s))\zeta(s, w) - f'(w(s))\zeta(s, w)\int b_x(w(s), y)u(s, dy) \\
&- b[w(s), u(s)]f''(w(s))\zeta(s, w) - \zeta(s, w)\int b_y(z, w(s))f'(z)u(s, dz)
\end{aligned}
$$

$$= \int \mu(dw')\{b_y(w(s), w'(s))f'(w(s))\zeta(s, w') \\ - b_y(w'(s), w(s))f'(w'(s))\zeta(s, w)\}.$$

But from the last expression of $G(s)$ it follows that $E_\mu\{G(s)\} = 0$ and so the lemma is proved.

We now write

$$\langle Y(t), f\rangle - \langle Y(0), f\rangle = \langle Y, f(w(t)) - f(w(0))\rangle \\ = \langle Y, N_f(t)\rangle + \int_0^t \langle Y(s), K_{u(s)}f\rangle ds + \int_0^t \langle Y(s), L_{u(s)}f\rangle ds .$$

If we set

$$B_f(t) = \langle Y, N_f(t)\rangle ,$$

then by (4.1), (2.14), Lemma 4.1 and (4.3) we have

$$E\{e^{i\{B_f(t)-B_f(s)\}} \mid \mathscr{F}_s\} = \exp\left\{-\int_s^t \|f'\|^2_{u(\tau)} d\tau/2\right\}, \quad 0 \le s \le t ,$$

where $\mathscr{F}_s$ is the smallest σ-field on Ω that makes $\{B_f(\tau), 0 \le \tau \le s, f \in \mathscr{S}\}$ measurable. Let $\{B(t)\}$ be a (distribution-valued) Brownian motion such that

$$\langle B(t), f\rangle = B_f(t) \quad \text{a.s.} \quad \text{(regularization)}.$$

Then

$$d\langle Y(t), f\rangle = d\langle B(t), f\rangle + \langle Y(t), K_{u(t)}f + L_{u(t)}f\rangle dt ,$$

or symbolically we have the infinite dimensional SDE

$$dY(t) = dB(t) + (K^*_{u(t)} + L^*_{u(t)})Y(t)dt ,$$

where $*$ means the dual operator. This is, of course, consistent with the following formal computation about the generator of the Markov process $Y(t)$.

Generators: $Y_n(t)$ is a (temporally inhomogeneous) Markov process and its generator $L_t^{(n)}$ "at time t" is given as follows. Let

$$\Phi(\eta) = \varphi(\langle \eta, f_1\rangle, \cdots, \langle \eta, f_m\rangle), \qquad f_1, \cdots, f_m \in \mathscr{S}, \quad \varphi \in C_0^\infty(\boldsymbol{R}^m).$$

Then for $\eta = \sqrt{n}(\bar{x} - u(t))$ ($\bar{x} = n^{-1}\sum \delta_{x_k}$) we have

$$(L_t^{(n)}\Phi)(\eta) = \sum_{\alpha,\beta=1}^m \left\langle \bar{x}, \frac{1}{2}f'_\alpha f'_\beta\right\rangle \partial^2_{\alpha\beta}\varphi + \sum_{\alpha=1}^m \langle \eta, K_{u(t)}f_\alpha\rangle \partial_\alpha\varphi \\ + \sum_{\alpha=1}^m \langle \bar{x}, b[\cdot, \eta]f'_\alpha\rangle \partial_\alpha\varphi ,$$

and hence the generator L_t of $Y(t)$ "at time t" is given (at least formally) by

$$(L_t\Phi)(\eta) = \lim_{n\to\infty} (L_t^{(n)}\Phi)(\eta) = \sum_{\alpha,\beta=1}^{m} \left\langle u(t), \frac{1}{2} f'_\alpha f'_\beta \right\rangle \partial^2_{\alpha\beta}\varphi$$
$$+ \sum_{\alpha=1}^{m} \langle \eta, K_{u(t)} f_\alpha \rangle \partial_\alpha \varphi + \sum_{\alpha=1}^{m} \langle \eta, L_{u(t)} f_\alpha \rangle \partial_\alpha \varphi .$$

§ 5. Large deviation

Let U_n be defined by (4) and let Q_n be its probability distribution, i.e.,

$$Q_n(A) = P\{U_n \in A\}, \quad A \subset \mathscr{P},$$

where $\mathscr{P}$ is the set of probability measures on W. For $\nu \in \mathscr{P}$ we set $v(t) = v(t, \cdot) = \nu\{w(t) \in \cdot\}$ and denote by $\tilde{\nu}$ the probability measure on W induced by the diffusion process (initial distribution $= u$) associated with the linear equation

$$\frac{\partial \tilde{u}}{\partial t} = \frac{1}{2} \cdot \frac{\partial^2 \tilde{u}}{\partial x^2} - \frac{\partial}{\partial x} \{b[x, v(t)]\tilde{u}\} .$$

We then define

$$(5.1) \qquad I_\mu(\nu) = \begin{cases} \int_W \log (d\nu/d\tilde{\nu}) d\nu , & \text{if } \nu \prec \mu \text{ and } \log (d\nu/d\tilde{\nu}) \in L^1(\nu) \\ \infty, & \text{otherwise.} \end{cases}$$

Theorem 5.1. (i) *For any closed set $C \subset \mathscr{P}$ and open set $G \subset \mathscr{P}$ (weak topology on $\mathscr{P}$)*

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log Q_n(C) \leq - \inf_{\nu \in C} I_\mu(\nu) ,$$

$$\underline{\lim_{n\to\infty}} \frac{1}{n} \log Q_n(G) \geq - \inf_{\nu \in G} I_\mu(\nu) .$$

(ii) *For any bounded weakly continuous function F on $\mathscr{P}$*

$$\lim_{n\to\infty} \frac{1}{n} \log E\{e^{nF(U_n)}\} = \sup_{\nu \in \mathscr{P}} \{F(\nu) - I_\mu(\nu)\} .$$

Proof. If we define a mapping $\theta: \mathscr{P} \to \mathscr{P}$ by $\theta(\rho) = \rho \circ \xi(\rho)^{-1}$, then $U_n = \theta(U_n^0)$ where $U_n^0 = n^{-1} \sum \delta_{w_k}$. Moreover, it is easy to see that θ is a homeomorphism from $\mathscr{P}$ onto itself and hence by a result of Donsker and Varadhan [3] for independent identically distributed random variables

$$\varlimsup_{n\to\infty} \frac{1}{n} \log Q_n(C) = \varlimsup_{n\to\infty} \log P\{U_n^0 \in \theta^{-1}(C)\}$$
$$\leq - \inf_{\rho \in \theta^{-1}(C)} I_\lambda^0(\rho),$$

where

$$I_\lambda^0(\rho) = \begin{cases} \int_W \log (d\rho/d\lambda)d\rho, & \text{if } \rho \prec \lambda \text{ and } \log (d\rho/d\lambda) \in L^1(\rho) \\ \infty, & \text{otherwise}. \end{cases}$$

We now set $\mu = \theta(\lambda)$ and $\nu = \theta(\rho)$. Since $\rho \prec \lambda \Leftrightarrow \nu \prec \bar{\nu} \equiv \lambda \circ \xi(\rho)^{-1} \Leftrightarrow \nu \prec \mu$ and $d\rho/d\lambda = g(\xi(w, \rho))$ where $g = d\nu/d\bar{\nu}$, we have

$$I_\lambda^0(\rho) = I_\mu(\nu), \qquad \inf_{\rho \in \theta^{-1}(C)} I_\lambda^0(\rho) = \inf_{\nu \in C} I_\mu(\nu),$$

proving the first inequality. The second inequality of (i) and the equality of (ii) can also be proved by the same method.

§ 6. The case of Boltzmann's equation

Kac's master equation approach to Boltzmann's equation goes back to 1956. The central limit theorem (fluctuation theory) in the time evolution scheme as in § 4 was discussed in [6], [10], [12], [13] and [15]. Although the methods in the preceding sections can not be applied in a straightforward way to the case of Boltzmann's equation, one may give some conjectures about fluctuations in the path space and large deviation problem for U_n. Here we consider McKean's 2-state model of Maxwellian molecules.

Let $\xi^{(n)}(t) = (\xi_1^{(n)}(t), \cdots, \xi_n^{(n)}(t))$ be a Markov process on the n-fold product space $\{\pm 1\}^n$ with generator

$$(K_n f)(x_1, \cdots, x_n)$$
$$= \frac{1}{n} \sum_{1\leq i<j\leq n} \sum{}' \{f(\cdots, x_i', \cdots, x_j', \cdots) - f(x_1, \cdots, x_n)\},$$

where $\sum'$ is the sum with respect to the two types of collisions

$$\begin{pmatrix} x_i' \\ x_j' \end{pmatrix} = \begin{pmatrix} x_i x_j \\ x_j \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} x_i \\ x_i x_j \end{pmatrix}.$$

Let u be a given probability measure on $\{\pm 1\}$ and assume that the initial distribution of $\xi^{(n)}(t)$ is $u \otimes \cdots \otimes u$. Sample paths of each component process $\xi_k^{(n)}(t)$ are in the space W of right continuous step functions taking values of ± 1 (the time is restricted to $0 \le t \le 1$ as before). Define U_n and Y_n as in (4) and (5). For each fixed $w \in W$ and $\varepsilon > 0$ we denote by $\mu_{w,\varepsilon}$ the probability measure on W induced by the Markov process associated with the nonlinear equation

$$\frac{d}{dt}\langle \tilde{u}(t), f\rangle = \langle \tilde{u}(t) \otimes (\tilde{u}(t) + \varepsilon\delta_{w(t)}), Kf\rangle \tag{6.1}$$

having initial distribution u, where $(Kf)(x, y) = f(xy) - f(x)$. $\mu = \mu_{w,0}$ is nothing but the probability law of the Markov process associated with the Boltzmann equation

$$\frac{d}{dt}\langle u(t), f\rangle = \langle u(t) \otimes u(t), Kf\rangle$$

of McKean's 2-state fictitious gas. We now define Q-functional by

$$Q(\varphi) = E_\mu\{|(I + \Lambda)\varphi|^2\} - |E_\mu\{(I + \Lambda)\varphi\}|^2$$

on the class of functions φ of the form

$$\varphi(w) = f(w(t_1), \cdots, w(t_m)), \quad f \in C_0^\infty(R^m), \quad 0 \le t_1 < \cdots < t_m,$$

where

$$\Lambda\varphi(w) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon}\{E_{\mu,w,\varepsilon}(\varphi) - E_\mu(\varphi)\},$$

and also define I-functional by

$$I_\mu(\nu) = \begin{cases} \log (d\nu/d\tilde{\nu})d\nu, & \text{if } \nu \prec \tilde{\nu} \text{ and } \log (d\nu/d\tilde{\nu}) \in L^1(\nu) \\ \infty, & \text{otherwise}, \end{cases}$$

where $\tilde{\nu}$ is the probability measure induced by the Markov process (initial distribution $= u$) associated with the linear equation

$$\frac{d}{dt}\langle \tilde{u}(t), f\rangle = \langle \tilde{u}(t) \otimes v(t), Kf\rangle,$$

$v(t)$ being defined by $v(t, \cdot) = \nu\{w(t) \in \cdot\}$. Then from the results in the diffusion case one may conjecture as follows.

$$\lim_{n\to\infty} E\{e^{iY_n(\varphi)}\} = e^{-Q(\varphi)/2}. \tag{6.2}$$

(6.3) The large deviation for U_n will be governed by the functional $I_\mu(\nu)$ as in Theorem 5.1.

Notes added in proof. Recently, making use of a method of Sznitman [11], T. Shiga and the author proved that (6.2) is correct.

References

[1] W. Braun and K. Hepp, The Vlasov dynamics and its fluctuations in the 1/N limit of interacting classical particles, Comm. Math. Phys., **56** (1977), 101–113.

[2] D. A. Dawson, Critical dynamics and fluctuations for a mean field model of cooperative behavior, J. Statistical Physics, **31** (1978), 29–85.

[3] M. D. Donsker and S. R. S. Varadhan, Asymptotic evaluation of certain Markov process expectations for large time–III, Comm. Pure Appl. Math., **29** (1976), 389–461.

[4] K. Itô, Motions of infinite particles (in Japanese), Kokyuroku RIMS, Kyoto Univ., **367** (1979), 1–33.

[5] M. Kac, Foundation of kinetic theory, Proceedings of 3rd Berkeley Symposium on Math. Stat. and Prob., vol. **3** (1956), 171–197.

[6] ——, Some probabilistic aspects of the Boltzmann equation, Acta Phys. Austriaca Supp. X, Springer Verlag, (1973), 379–400.

[7] S. Kusuoka and Y. Tamura, The convergence of Gibbs measures associated with mean field potentials, to appear in J. Fac. Sci., Univ. of Tokyo.

[8] H. P. McKean, A class of Markov processes associated with nonlinear parabolic equations, Proc. Nat. Acad. Sci., **56** (1966), 1907–1911.

[9] ——, Propagation of chaos for a class of non-linear parabolic equations, Lecture Series in Differential Equations **7**, Catholic Univ. (1967), 41–57.

[10] ——, Fluctuations in the kinetic theory of gases, Comm. Pure Appl. Math., **28** (1975), 435–455.

[11] A. S. Sznitman, An example of non-linear diffusion process with normal reflecting boundary condition and some related limit theorems, preprints, 1983.

[12] H. Tanaka, Fluctuation theory for Kac's one dimensional model of Maxwellian molecules, Sankhyā: The Indian Journal of Statistics, **44** (1982), Series A. Pt. I, 23–26.

[13] ——, Some probabilistic problems in the spatially homogeneous Boltzmann equation, Proc. IFIP-WG 7/1 Working Conference on Theory and Applications of Random Fields held in Bangalore, January 1982, Lecture Notes in Control and Information Sciences **49**, Springer-Verlag.

[14] H. Tanaka and M. Hitsuda, Central limit theorem for a simple diffusion model of interacting particles, Hiroshima Math. J., **11** (1981), 415–423.

[15] K. Uchiyama, Fluctuations of Markovian systems in Kac's caricature of a Maxwellian gas, J. Math. Soc. Japan, **35** (1983), 477–499.

DEPARTMENT OF MATHEMATICS
FACULTY OF SCIENCE AND TECHNOLOGY
KEIO UNIVERSITY,
YOKOHAMA 223, JAPAN

Z. Wahrscheinlichkeitstheorie verw. Gebiete
69, 439-459 (1985)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Central Limit Theorem for a System of Markovian Particles with Mean Field Interactions

Tokuzo Shiga[1] and Hiroshi Tanaka[2]

[1] Department of Applied Physics, Faculty of Science, Tokyo Institute of Technology, Oh-okayama, Tokyo, Japan

[2] Department of Mathematics, Faculty of Science and Technology, Keio University, Yokohama, Japan

§ 0. Introduction

Let S be a measurable space and $\mathscr{P}(S)$ be the totality of probability measures on S. Suppose that to each $v \in \mathscr{P}(S)$ there corresponds a generator Q_v of a Markov process on S. Starting from the family we can consider an interacting n-particle system $X^{(n)}(t)=(X_1^{(n)}(t), \ldots, X_n^{(n)}(t))$, which is a Markov process on $S^{\otimes n}=S \times \ldots \times S$ with generator

(0.1) $$Q^{(n)} \phi(x_1, \ldots, x_n)=\sum_{i=1}^{n} Q_{v_n}^{(i)} \phi(x_1, \ldots, x_n), \quad v_n=\frac{1}{n} \sum_{j=1}^{n} \delta_{x_j},$$

where δ_x stands for the δ-distribution at x, and $Q_v^{(i)}$ is used instead of Q_v when it acts on the i-th variable of $\phi(x_1, \ldots, x_n)$.

Denoting by W the path space on S, let us consider an empirical distribution U_n of the n-particle system:

(0.2) $$U_n=\frac{1}{n} \sum_{i=1}^{n} \delta_{X_i^{(n)}(\cdot)},$$

which is a random measure on W.

Our problem is described as follows. Assume that $(X_1^{(n)}(0), \ldots, X_n^{(n)}(0))$ is independent and identically distributed with common distribution $u \in \mathscr{P}(S)$ for any $n \geqq 1$. It is known that there exists a probability distribution μ on the path space W such that U_n converges to μ as $n \to \infty$ in the sense that $\langle U_n, \Phi \rangle$ converges to $\langle \mu, \Phi \rangle$ in probability for any bounded measurable function Φ on W (the law of large numbers). Then the marginal distribution $u(t)$ of μ solves the following non-linear equation

(0.3) $$\frac{d}{dt} \langle u(t), \phi \rangle=\langle u(t), Q_{u(t)} \phi \rangle, \quad (\phi\text{: test function}).$$

As the next stage we consider the central limit theorem on the path space, that is to show that

$$\xi_n=\sqrt{n}(U_n-\mu) \tag{0.4}$$

converges as $n\to\infty$ to a Gaussian field.

In the case of diffusion processes there have been many works concerning these problems (cf. Dawson [1], Kusuoka-Tamura [7], McKean [9], Sznitman [13], [14], Tanaka [15], Tanaka-Hitsuda [16]). Above all, Sznitman recently proved the central limit theorem on the path space for a general class of diffusion processes on $\boldsymbol{R}^d$ associated with $\{Q_v;\ v\in\mathscr{P}(\boldsymbol{R}^d)\}$ where

$$Q_v=\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}a_{ij}[x,v]\frac{\partial^2}{\partial x_i\partial x_j}+\sum_{i=1}^{d}b_i[x,v]\frac{\partial}{\partial x_i}, \tag{0.5}$$

$\{\alpha_{ij}(x,y)\}_{1\leq i,j\leq d}$ and $\{b_i(x,y)\}_{1\leq i\leq d}$ being uniformly bounded and Lipschitz continuous functions on $\boldsymbol{R}^d\times\boldsymbol{R}^d$ and

$$a_{ij}[x,v]=\sum_{k=1}^{d}\int\alpha_{ik}(x,y)\,v(dy)\int\alpha_{jk}(x,y)\,v(dy),$$

$$b_i[x,v]=\int b_i(x,y)\,v(dy).$$

In the present paper we study the central limit theorem mainly for pure jump type Markov processes. More precisely, let $(S,\mathscr{B}_S)$ be a measurable space and let $Q(x,x';dy)$ be a bounded measurable kernel on $S\times S\times\mathscr{B}_S$. Then to each $v\in\mathscr{P}(S)$ there corresponds a generator Q_v of a Markov process on S by

$$Q_v\phi(x)=\int v(dx')\int Q(x,x';dy)(\phi(y)-\phi(x)). \tag{0.6}$$

Under a certain condition of non-degeneracy we obtain the central limit theorem for (0.4) (see § 2).

Furthermore, we remark in § 3 that our proof is valid for McKean's 2-state model of Boltzmann's equation with a slight modification and we have the central limit theorem as conjectured in [15]. We also give some remarks in § 4 to the case of diffusion processes as discussed in [13], [14] and [15]; here we assume the diffusion matrix $\alpha(x,y)$ is the identity matrix but the drift vector $b(x,y)$ is bounded measurable.

Our method of the proof is essentially based on Sznitman's paper [13], which treats the case of diffusion processes with $\alpha(x,y)=$ identity matrix. There he calculates a limiting quantities of Cameron-Martin-Maruyama-Girsanov density by making use of symmetric statistics and multiple Wiener integrals. His method is also applicable to our case. However we emphasize that once we obtain formulae on multiple Wiener integrals (the two lemmas in § 1) the proof turns out to be more transparent.

§ 1. Symmetric Statistics and Multiple Wiener Integrals

For the later use we summarize some facts on symmetric statistics and multiple Wiener integrals according to Dynkin and Mandelbaum [3], and present two formulae on multiple Wiener integrals.

Let $(\mathfrak{X}, \mathscr{B})$ be a separable Borel space, and let $\{X_n\}_{n=1}^{\infty}$ be a sequence of independent and identically distributed $\mathfrak{X}$-valued random variables with a distribution ν. For each $k=1,2,\dots$ let $L^2(\nu^{\otimes k})(L_c^2(\nu^{\otimes k}))$ be the space of all real-valued (complex-valued) square integrable functions on $\mathfrak{X}^{\otimes k}=\mathfrak{X}\times\dots\times\mathfrak{X}$ (k-fold) with respect to $\nu^{\otimes k}=\nu\times\dots\times\nu$ (k-fold). Denote by $L_{\mathrm{symm}}^2(\nu^{\otimes k})$ the space of all symmetric functions of $L^2(\nu^{\otimes k})$. Then all of these spaces are separable Hilbert spaces. To $h_k\in L_{\mathrm{symm}}^2(\nu^{\otimes k})$ there corresponds a symmetric statistic $\sigma_k^n(h_k)$ defined by

$$
\begin{aligned}
(1.1)\qquad \sigma_k^n(h_k) &= \sum_{1\leqq i_1<\dots<i_k\leqq n} h_k(X_{i_1},\dots,X_{i_k}) && \text{for } n\geqq k\\
&=0 && \text{for } n<k.
\end{aligned}
$$

Let $\{I_1(\phi);\phi\in L^2(\nu)\}$ be a centered Gaussian field satisfying

$$(1.2)\qquad E\{I_1(\phi)I_1(\psi)\}=(\phi,\psi)_{L^2(\nu)}.$$

For each $\phi\in L^2(\nu)$ set

$$(1.3)\qquad h_0^\phi=1,\quad h_k^\phi(x_1,\dots,x_k)=\phi(x_1)\dots\phi(x_k)\quad k=1,2,\dots.$$

A multiple Wiener integral $I_k(h_k^\phi)$ is defined by the following relation,

$$(1.4)\qquad \sum_{k=0}^{\infty}\frac{t^k}{k!}I_k(h_k^\phi)=\exp\left(tI_1(\phi)-\frac{t^2}{2}\|\phi\|_{L^2(\nu)}^2\right).$$

For a general $h_k\in L_{\mathrm{symm}}^2(\nu^{\otimes k})$ $I_k(h_k)$ is defined by a standard procedure making use of the fact that the linear hull of $\{h_k^\phi:\phi\in L^2(\nu)\}$ is dense in $L_{\mathrm{symm}}^2(\nu^{\otimes k})$. Then we have

Theorem 1.1 (Dynkin-Mandelbaum [3]). *Suppose $\{h_k\}_{k=1}^{\infty}$ satisfies the following: for each $k\geqq 1$ $h_k\in L_{\mathrm{symm}}^2(\nu^{\otimes k})$, and*

$$(1.5)\qquad \int h_k(x_1,\dots,x_{k-1},x)\,\nu(dx)=0\quad \textit{for } \nu^{\otimes k-1}\text{-a.e.}\,(x_1,\dots,x_{k-1}).$$

Then

$$\{n^{-\frac{k}{2}}\sigma_k^n(h_k);k\geqq 1\}\ \textit{converges to}\ \left\{\frac{1}{k!}I_k(h_k);k\geqq 1\right\}\quad \textit{as } n\to\infty$$

in the sense of convergence of finite dimensional distributions.

Let $a(x,y)\in L^2(\nu\otimes\nu)$ and denote by A the integral operator on $L_c^2(\nu)$ associated with $a(x,y)$,

$$(1.6)\qquad A\phi(x)=\int a(x,y)\,\phi(y)\,\nu(dy)\qquad (\phi\in L_c^2(\nu)).$$

Then A is a Hilbert-Schmidt operator, and $A^n\,(n\geqq 2)$ and AA^* are trace class operators. It is easy to see

$$(1.7)\qquad \operatorname{Trace}A^n=\int\dots\int a(x_1,x_2)\,a(x_2,x_3)\dots a(x_n,x_1)\,\nu(dx_1)\dots\nu(dx_n)$$

for any $n \geqq 2$, and

$$\text{Trace}\, A A^* = \|a\|^2_{L^2(\nu \otimes \nu)}. \tag{1.8}$$

The following lemmas will be used effectively in proving central limit theorems in the following sections.

Lemma 1.2. *Suppose that* Trace $A^n = 0$ *for all* $n \geqq 2$. *Then*

$$E[e^{\frac{1}{2} I_2(f)}] = e^{\frac{1}{2} \text{Trace}\, A A^*} \tag{1.9}$$

where

$$f(x, y) = a(x, y) + a(y, x) - \int a(x, z)\, a(y, z)\, \nu(dz). \tag{1.10}$$

Proof. Theorem 9.3 of Simon [12] implies that Trace $A^n = 0$ for all $n \geqq 2$ if and only if

$$\det_2(I + \mu A) = 1 \quad \text{for any complex number } \mu, \tag{1.11}$$

where $\det_2(I + \mu A)$ stands for the regularized determinant. In particular we have

$$\det_2(I - A) = \det_2(I - A^*) = 1. \tag{1.12}$$

Hence it follows from Theorem 9.2 of [12] that $I - A^*$ is invertible. Denoting by F the integral operator on $L^2_c(\nu)$ associated with $f(x, y)$, $I - F = (I - A)(I - A^*)$ is strictly positive definite because of the invertibility of $I - A^*$. Accordingly any eigen-value of F is real and less than one.

Let $\lambda_1, \lambda_2, \ldots$ be all eigen-values of F and $\{e_1, e_2, \ldots\}$ be the orthonormal base of the corresponding eigen-vectors of F on $L^2(\nu)$. Then f is represented as follows:

$$f = \sum_{k=1}^{\infty} \lambda_k e_k \otimes e_k.$$

Since $\sum_{k=1}^{\infty} \lambda_k^2 < \infty$ and $\lambda_k < 1$ for any $k \geqq 1$, we get using the relation (1.4) and the independence of $\{I_1(e_k)\}_{k=1}^{\infty}$

$$\begin{aligned} E[e^{\frac{1}{2} I_2(f)}] &= E\left[\exp \sum_{k=1}^{\infty} \frac{\lambda_k}{2} (I_1(e_k)^2 - 1)\right] \\ &= \prod_{n=1}^{\infty} E\left[\exp \frac{\lambda_k}{2} (I_1(e_k)^2 - 1)\right] \\ &= \prod_{k=1}^{\infty} \frac{e^{-\frac{\lambda_k}{2}}}{\sqrt{1 - \lambda_k}}. \end{aligned} \tag{1.13}$$

On the other hand the following formulae are known (cf. Simon [12], p. 107):

$$\det_2(I - F) = \prod_{k=1}^{\infty} (1 - \lambda_k) e^{\lambda_k} \tag{1.14}$$

$$\det_2(I - A)(I - A^*) = \det_2(I - A) \det_2(I - A^*) e^{-\text{Trace}\, A A^*}. \tag{1.15}$$

Thus we have by (1.12)

$$\det_2(I-F)=\det_2(I-A)(I-A^*)=e^{-\operatorname{Trace} AA^*}. \tag{1.16}$$

Therefore (1.9) follows immediately from (1.13)–(1.16).

Lemma 1.3. *Assume the same condition as in Lemma 1.2. Then $I-A$ is invertible and for any $\phi \in L^2(v)$*

$$\begin{aligned} &E[\exp\{\sqrt{-1}\,I_1(\phi)+\tfrac{1}{2}I_2(f)\}] \\ &\quad =\exp -\tfrac{1}{2}\{\|(I-A)^{-1}\phi\|^2_{L^2(v)}-\operatorname{Trace} AA^*\}. \end{aligned} \tag{1.17}$$

Proof. It follows from (1.12) that $I-A$ is invertible. Also using the same $\{e_k\}$ of the proof of Lemma 1.2 we have

$$\begin{aligned} &E[\exp\{\sqrt{-1}\,I_1(\phi)+\tfrac{1}{2}I_2(f)\}] \\ &=\prod_{k=1}^{\infty} E\left[\exp\left\{\sqrt{-1}(\phi,e_k)I_1(e_k)+\frac{\lambda_k}{2}(I_1(e_k)^2-1)\right\}\right] \\ &=\prod_{k=1}^{\infty}\frac{e^{-\frac{\lambda_k}{2}}}{\sqrt{1-\lambda_k}}\exp\left\{-\frac{1}{2}\sum_{k=1}^{\infty}(1-\lambda_k)^{-1}(\phi,e_k)^2\right\} \\ &=\{\det_2(I-F)\}^{-\frac{1}{2}}\exp\{-\tfrac{1}{2}((I-F)^{-1}\phi,\phi)_{L^2(v)}\} \\ &=e^{\frac{1}{2}\operatorname{Trace} AA^*}\exp\{-\tfrac{1}{2}\|(I-A)^{-1}\phi\|^2_{L^2(v)}\}. \end{aligned} \tag{1.18}$$

§2. The Case of Pure Jump Type Markov Processes

Let S be a standard Borel space, i.e., $(S,\mathscr{B}_S)$ is Borel isomorphic with a complete separable metric space. Denote by $\mathbb{B}(S)$ the Banach space of all bounded measurable functions on S equipped with the supremum norm $\|\cdot\|_\infty$.

Let $Q(x,x';dy)$ be a measure kernel satisfying the following conditions:

(2.1) for any fixed $(x,x')\in S\times S$ $Q(x,x';dy)$ is a bounded measure on $(S,\mathscr{B}_S)$ and $Q(x,x';\{x\})=0$;

(2.2) for any fixed $E\in\mathscr{B}_S$ $Q(x,x';E)$ is measurable in (x,x') with respect to $\mathscr{B}_S \otimes\mathscr{B}_S$;

(2.3) $q(x,x')=Q(x,x';S\setminus\{x\})$ is bounded in $(x,x')\in S\times S$;

(2.4) there exist constants $c_1>0$ and $c_2>0$ such that $c_1 Q(x,x';dy)\leq Q(x,x'';dy)\leq c_2 Q(x,x';dy)$ for any x, x' and x'' of S.

For each finite measure u on S, we write

$$Q_u(x;dy)=\int_S Q(x,x';dy)\,u(dx'), \tag{2.5}$$

and define a bounded operator Q_u on $\mathbb{B}(S)$ by

$$Q_u\phi(x)=\int_S Q_u(x;dy)(\phi(y)-\phi(x)) \quad \text{for } \phi\in\mathbb{B}(S). \tag{2.6}$$

Let W be the set of all S-valued step functions, i.e. for each $w\in W$ there exist $0=t_1<t_2<\ldots<t_n<\ldots$ with $t_n\to\infty$ such that

$$w(t)=w(t_i) \quad \text{for } t_i\leqq t<t_{i+1}. \tag{2.7}$$

Denote by $\mathscr{F}$ and $\mathscr{F}_t$ the usual σ-field of W. For any $T>0$ $(W_T,\mathscr{F}_T)$ denotes the restriction of $(W,\mathscr{F})$ to $[0,T]$. Then these are separable Borel spaces. Let $\mu\in\mathscr{P}(W)$, the set of probability measures on $(W,\mathscr{F})$. μ is called a McKean measure (or a McKean process) corresponding to $\{Q_v; v\in\mathscr{P}(S)\}$ if

$$\phi(w(t))-\int_0^t Q_{u(s)}\phi(w(s))\,ds \quad \text{is a } \mu\text{-martingale for any } \phi\in\mathbb{B}(S), \tag{2.8}$$

$$\mathscr{L}(w(t);\mu)=u(t). \tag{2.9}$$

Here $\mathscr{L}(w(t);\mu)$ stands for the probability distribution of $w(t)$ under μ.

For any fixed $u\in\mathscr{P}(S)$, the set of probability measures on $(S,\mathscr{B}_S)$, there exists a unique McKean measure μ corresponding to $\{Q_v; v\in\mathscr{P}(S)\}$ with $\mathscr{L}(w(0);\mu)=u$, which will be shown in Lemma 2.3.

Now, let us consider an n-particle system, which is a Markov process $(X_1^n(t), \ldots, X_n^n(t))$ on $S^{\otimes n}=S\times\ldots\times S$ generated by the following operator:

$$Q^{(n)}\phi(x_1,\ldots,x_n)=\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\int_S Q(x_i,x_j;dy)\Delta_i^y\phi(x_1,\ldots,x_n) \tag{2.10}$$

where $\Delta_i^y\phi(x_1,\ldots,x_n)=\phi(x_1,\ldots,\underset{\text{the } i\text{-th}}{y},\ldots,x_n)-\phi(x_1,\ldots,x_n)$.

We assume the initial distribution is $u^{\otimes n}=u\otimes\ldots\otimes u$, the n-fold product measure of u. Let us introduce an empirical distribution U_n of the n-particle system,

$$U_n=\frac{1}{n}\sum_{i=1}^{n}\delta_{X_i^n(\cdot)}. \tag{2.11}$$

Then U_n is a random variable with values in $\mathscr{P}(W)$, and we can show that U_n converges as $n\to\infty$ to the McKean measure μ.

In order to discuss a central limit theorem for U_n let us introduce some notations. From now on, we fix $T>0$. For any $w\in W_T$, $0\leqq t\leqq T$ and $E\in\mathscr{B}_S$ set

$$N(t,w;E)=\sum_{s\leqq t} I(w(s)\in E, w(s)\neq w(s-)), \tag{2.12}$$

and

$$\tilde{N}(t,w;E)=N(t,w,E)-\int_0^t Q_{u(s)}(w(s);E)\,ds, \tag{2.13}$$

where $I(B)$ stands for the indicator function of B. For each $(w,w')\in W_T\times W_T$ set

$$a(w,w')=\int_S\int_{[0,T]}(q_t(w'(s-),w(s-);y)-1)\,\tilde{N}(ds,dy;w') \tag{2.14}$$

where

$$q_t(x,x';y)=Q(x,x';dy)/Q_{u(t)}(x;dy),$$

$$u(t)=\mathscr{L}(w(t);\mu).$$

We note $q_t(x,x';y)$ is bounded by (2.4). Furthermore since $a(w,w')$ is square integrable with respect to $\mu\otimes\mu$, there corresponds a Hilbert-Schmidt operator A on $L_c^2(W_T,\mu)$,

$$A\phi(w)=\int_{W_T} a(w,w')\,\phi(w')\,\mu(dw'). \tag{2.15}$$

Note that $L_c^2(W_T,\mu)$ is a separable complex Hilbert space. Then it is shown that $I-A$ is an invertible operator.

Let us describe our main results. Let

$$\xi_n(\Phi)=\sqrt{n}\,\langle U_n-\mu,\Phi\rangle=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\Phi}(w_i)\quad\text{for } \Phi\in L^2(W_T,\mu) \tag{2.16}$$

where $\tilde{\Phi}=\Phi-\langle\mu,\Phi\rangle$.

Theorem 2.1. *$\{\xi_n(\Phi)\colon\Phi\in L^2(W_T,\mu)\}$ converges as $n\to\infty$ to a centered Gaussian field $\{\xi(\Phi)\colon\Phi\in L^2(W_T,\mu)\}$ satisfying*

$$E[\xi(\Phi)\xi(\Psi)]=((I-A)^{-1}\tilde{\Phi},(I-A)^{-1}\tilde{\Psi})_{L^2(W_T,\mu)} \tag{2.17}$$

in the sense of convergence of finite dimensional distributions.

Next let us consider another interpretation of the operator $(I-A)^{-1}$ as in Tanaka [15] and Sznitman [13]. For any fixed $\zeta\in W$ and $\varepsilon>0$ there exists a unique $\mu_\varepsilon^\zeta\in\mathscr{P}(W)$ such that

$$\phi(w(t))-\int_0^t Q_{(u_\varepsilon^\zeta(s)+\varepsilon\delta_{\zeta(s)})}\phi(w(s))\,ds \tag{2.18}$$

is a μ_ε^ζ-martingale for any $\phi\in\mathbb{B}(S)$, and

$$\mathscr{L}(w(t);\mu_\varepsilon^\zeta)=u_\varepsilon^\zeta(t),\qquad \mathscr{L}(w(0);\mu_\varepsilon^\zeta)=u. \tag{2.19}$$

For any $\Phi\in L^2(W_T,\mu)$ and $\varepsilon>0$

$$\Pi_\varepsilon\Phi(\zeta)=\frac{1}{\varepsilon}\langle\mu_\varepsilon^\zeta-\mu,\Phi\rangle \tag{2.20}$$

is well-defined, and we can show

$$\Pi\Phi=\text{s-}\lim_{\varepsilon\to 0}\Pi_\varepsilon\Phi\qquad\text{exists in } L^2(W_T,\mu). \tag{2.21}$$

Furthermore the following identity holds.

Theorem 2.2

$$I+\Pi=(I-A)^{-1}. \tag{2.22}$$

We will prove these theorems by a series of lemmas.

Lemma 2.3. *Assume (2.1)-(2.3). Then for any $u\in\mathscr{P}(S)$ there exists a unique McKean measure μ corresponding to $\{Q_v;v\in\mathscr{P}(S)\}$ with $\mathscr{L}(w(0);\mu)=u$.*

Proof. Let $v(t)$ be a $\mathscr{P}(S)$-valued function defined on $[0, \infty)$ such that $\langle v(t), \phi\rangle$ is Borel measurable in t for any $\phi \in \mathbb{B}(S)$. We first claim that there exists a unique solution $\nu \in \mathscr{P}(W)$ of the following martingale problem:

$$\phi(w(t)) - \int_0^t Q_{v(s)}\phi(w(s))\,ds \tag{2.23}$$

is a ν*-martingale for any* $\phi \in \mathbb{B}(S)$, *and*

$$\mathscr{L}(w(0); \nu) = v(0). \tag{2.24}$$

For a $c > \|q\|_\infty = \sup_{x,x'} q(x, x')$ define a probability kernel P_t by

$$P_t(x; E) = \frac{1}{c} Q_{v(t)}(x; E \setminus \{x\}) + \left(1 - \frac{q[x, v(t)]}{c}\right) \delta_{\{x\}}(E) \tag{2.25}$$

where $x \in S$, $E \in \mathscr{B}_S$ and $q[x, v] = \int_S q(x, y)\, v(dy)$. Then

$$Q_{v(t)}\phi(x) = c\left[\int P_t(x; dy)\,\phi(y) - \phi(x)\right]. \tag{2.26}$$

Let $X(0)$ be an S-valued randon variable with the distribution $v(0)$, and N_t be a Poisson process on $\{0, 1, 2, \ldots\}$ with the intensity c independent of $X(0)$. Denote by σ_n the n-th jumping time of N_t. Let us construct an S-valued process $X(t)$. Set $X(t) = X(0)$ for $0 \leqq t < \sigma_1$. At $t = \sigma_1$, $X(\sigma_1)$ is a randomly chosen point of S according to the transition law $P_{\sigma_1}(X(\sigma_1 -); dy)$, and set $X(t) = X(\sigma_1)$ for $\sigma_1 \leqq t < \sigma_2$. Repeating this procedure we get an S-valued process $X(t)$. It is a routine work to see the distribution μ on W induced by $X(t)$ is a unique solution of the above martingale problem (2.23) and (2.24).

Next we note for any $u \in \mathscr{P}(S)$ the following equation has a unique $\mathscr{P}(S)$-valued solution

$$\langle u(t), \phi\rangle - \langle u, \phi\rangle = \int_0^t \langle u(s), Q_{u(s)}\phi\rangle\, ds \quad \textit{for any } \phi \in \mathbb{B}(S). \tag{2.27}$$

Let $u^0(t) = u$. If $u^{n-1}(t)$ is defined we have a unique solution μ^n of (2.23) and (2.24) with $v(t) = u^{n-1}(t)$, and set $u^n(t) = \mathscr{L}(w(t); \mu^n)$. Then we have

$$\langle u^n(t), \phi\rangle - \langle u, \phi\rangle = \int_0^t \langle u^n(s), Q_{u^{n-1}(s)}\phi\rangle\, ds \tag{2.28}$$

for any $\phi \in \mathbb{B}(S)$. Noting (2.3) it follows

$$\begin{aligned} &\|u^{n+1}(t) - u^n(t)\|_{\mathrm{var}} \\ &\leqq 2\|q\|_\infty \int_0^t (\|u^{n+1}(s) - u^n(s)\|_{\mathrm{var}} + \|u^n(s) - u^{n-1}(s)\|_{\mathrm{var}})\, ds, \end{aligned} \tag{2.29}$$

where $\|\cdot\|_{\mathrm{var}}$ stands for the total variation norm. This yields that there exists $u(t) \in \mathscr{P}(S)$ such that

$$\lim_{n \to \infty} \sup_{0 \leqq t \leqq T} \|u^n(t) - u(t)\|_{\mathrm{var}} = 0 \quad \textit{for any } T > 0. \tag{2.30}$$

Thus we get a solution $u(t)$ of (2.27), and the uniqueness also is obvious.

Finally, let μ be a unique solution of (2.23) with $v(t)=u(t)$. Then $\tilde{u}(t) = \mathscr{L}(w(t);\mu)$ is a solution of

$$\langle \tilde{u}(t), \phi\rangle - \langle u, \phi\rangle = \int_0^t \langle u(s), Q_{u(s)}\phi\rangle\, ds \qquad \phi \in \mathbb{B}(S). \tag{2.31}$$

Since the uniqueness of solutions to the linear equation (2.31) can also be proved easily, $\tilde{u}(t)$ coincides with $u(t)$ which is also a solution of (2.27). Therefore μ is a McKean measure corresponding to $\{Q_v; v\in\mathscr{P}(S)\}$ with $\mathscr{L}(w(0);\mu)=u$ which is also unique by virtue of the uniqueness result for (2.27) and (2.23).

From now on, we fix the McKean measure μ obtained in Lemma 2.3.

Lemma 2.4 (i). *For any* $E\in\mathscr{B}_S$

$$\tilde{N}(t,E;w) \text{ is a } \mu\text{-martingale, and} \tag{2.32}$$

(ii) *for any* E_1 *and* E_2 *of* $\mathscr{B}_S$

$$\tilde{N}(t,E_1;w)\,\tilde{N}(t,E_2;w) - \int_0^t Q_{u(s)}(w(s); E_1\cap E_2)\, ds \tag{2.33}$$

is a μ*-martingale.*

Proof. Noting that for any $\phi\in\mathbb{B}(S)$

$$\phi(w(t)) - \phi(w(0)) = \int_{[0,t]\times S} (\phi(y) - \phi(w(s-)))\, N(ds, dy; w), \quad \text{and} \tag{2.34}$$

$$\phi(w(t)) - \int_0^t Q_{u(s)}\phi(w(s))\, ds \qquad \text{is a } \mu\text{-martingale,} \tag{2.35}$$

we have

$$\int_{[0,t]\times S} (\phi(y) - \phi(w(s-)))\, \tilde{N}(ds, dy; w) \qquad \text{is a } \mu\text{-martingale.} \tag{2.36}$$

Also it is easily seen that for $\psi\in\mathbb{B}(S)$

$$\int_{[0,t]\times S} \psi(w(s-))(\phi(y) - \phi(w(s-)))\, \tilde{N}(ds, dy; w) \qquad \text{is a } \mu\text{-martingale.} \tag{2.37}$$

In particular, if $\phi(x)\psi(x)=0$ for all $x\in S$ then

$$\int_{[0,t]\times S} \phi(y)\psi(w(s-))\, \tilde{N}(ds, dy; w) \qquad \text{is a } \mu\text{-martingale.} \tag{2.38}$$

Furthermore noticing $\tilde{N}((s,y); w(s-)=y,\ 0\leqq s\leqq t)=0$ for any $t>0$ it is easy to see that for any $f\in\mathbb{B}(S\times S)$

$$\int_{[0,t]\times S} f(y, w(s-))\, \tilde{N}(ds, dy; w) \qquad \text{is a } \mu\text{-martingale,} \tag{2.39}$$

from which (i) follows. (ii) is immediate from (i) combined with the following identity obtained by integration by parts,

$$\tilde N(t,E_1;w)\tilde N(t,E_2;w)-N(t,E_1\cap E_2;w) = \int_0^t \tilde N(s-,E_1;w)\,d\tilde N(s,E_2;w)+\int_0^t \tilde N(s-,E_2;w)\,d\tilde N(s,E_1;w). \tag{2.40}$$

Let $(X^n(t)=(X_1^n(t),\dots,X_n^n(t)))$ be the Markov process on $S^{\otimes n}$ with the initial distribution $u^{\otimes n}$, generated by $Q^{(n)}$ of (2.10). Let us denote by P^n the probability distribution induced by $\{X^{(n)}(t)\}_{0\le t\le T}$ on the path space $(W_T^n,\mathscr{F}_T^n)$, where $W_T^n=W_T\times\dots\times W_T$ and $\mathscr{F}_t^n=\mathscr{F}_t\times\dots\times\mathscr{F}_t$ for $0\le t\le T$.

Lemma 2.5. *For each $n\ge 1$ P^n and $\mu^{\otimes n}$ are mutually absolutely continuous and*

$$\frac{dP^n}{d\mu^{\otimes n}}(\mathbf{w})=\exp H_T^n(\mathbf{w}), \tag{2.41}$$

$$H_t^n(\mathbf{w})=\sum_{i=1}^n \int_{[0,t]\times S}\left\{\log\frac{1}{n}\sum_{j=1}^n q_s(w_i(s-),w_j(s-);y)\right\}N(ds,dy;w_i) -\int_0^t\left(\frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n q(w_i(s),w_j(s))-\sum_{i=1}^n q[w_i(s),u(s)]\right)ds \tag{2.42}$$

for $\mathbf{w}=(w_1,\dots,w_n)\in W_T^n$ and $t>0$.

Proof. Let $M_t^n(\mathbf{w})=\exp H_t^n(\mathbf{w})$. Since $\{w_i(t)\}_{1\le i\le n}$ have no common jumps almost surely $(\mu^{\otimes n})$, it follows from Itô's formula (cf. Ikeda-Watanabe [4], p. 66)

$$M_t^n(\mathbf{w})-1=\sum_{i=1}^n\int_{]0,t]\times S} M_{s-}^n\left(\frac{1}{n}\sum_{j=1}^n q_s(w_i(s-),w_j(s-);y)-1\right)\tilde N(ds,dy;w_i). \tag{2.43}$$

Hence $M_t^n(\mathbf{w})$ is a $\mu^{\otimes n}$-martingale.

Next, for $\phi\in\mathbb{B}(S^{\otimes n})$, set

$$K_\phi(t)=\phi(\mathbf{w}(t))-\int_0^t Q^{(n)}\phi(\mathbf{w}(s))\,ds \tag{2.44}$$

$$\tilde K_\phi(t)=\phi(\mathbf{w}(t))-\sum_{i=1}^n\int_0^t\int_S Q_{u(s)}(w_i(s);dy)\,\Delta_i^y\phi(\mathbf{w}(s))\,ds. \tag{2.45}$$

It is obvious that $\tilde K_\phi(t)$ is a $\mu^{\otimes n}$-martingale. We will show

$$M_t^n K_\phi(t)\quad \textit{is also a } \mu^{\otimes n}\textit{-martingale.} \tag{2.46}$$

By integration by parts

$$M_t^n K_\phi(t)=\int_0^t K_\phi(s-)\,dM_s^n+\int_0^t M_{s-}^n\,d\tilde K_\phi(s) +\int_0^t M_{s-}^n(dK_\phi(s)-d\tilde K_\phi(s))+\sum_{s\le t}(M_s^n-M_{s-}^n)(K_\phi(s)-K_\phi(s-)). \tag{2.47}$$

Also,

(2.48)
$$\begin{aligned}&\sum_{s\leqq t}(M_s^n-M_{s-}^n)(K_\phi(s)-K_\phi(s-))\\&=\sum_{s\leqq t}(M_s^n-M_{s-}^n)(\phi(\mathbf{w}(s))-\phi(\mathbf{w}(s-)))\\&=\sum_{i=1}^n\int M_{s-}^n\left(\frac{M_s^n}{M_{s-}^n}-1\right)\Delta_i^y\phi(\mathbf{w}(s-))N(ds,dy;w_i)\\&=\sum_{i=1}^n\int_{[0,t]\times S}M_{s-}^n\left(\frac{1}{n}\sum_{j=1}^n q_s(w_i(s-),w_j(s-);y)-1\right)\\&\quad\cdot\Delta_i^y\phi(\mathbf{w}(s))N(ds,dy;w_i).\end{aligned}$$

On the other hand noting that

(2.49)
$$\begin{aligned}&\int_0^t M_{s-}^n(dK_\phi(s)-d\tilde{K}_\phi(s))\\&=-\sum_{i=1}^n\int_0^t\int_S M_{s-}^n\left(\frac{1}{n}\sum_{j=1}^n q_s(w_i(s),w_j(s);y)-1\right)Q_{u(s)}(w_i(s);dy)\Delta_i^y\phi(\mathbf{w}(s))\,ds,\end{aligned}$$

we have

(2.50)
$$\begin{aligned}M_t^nK_\phi(t)=&\int_0^t K_\phi(s-)\,dM_s^n+\int_0^t M_{s-}^n\,d\tilde{K}_\phi(s)\\&+\sum_{i=1}^n\int_{[0,t]\times S}M_{s-}^n\left(\frac{1}{n}\sum_{j=1}^n q_s(w_i(s-),w_j(s-);y)-1\right)\\&\cdot\Delta_i^y\phi(\mathbf{w}(s))\tilde{N}(ds,dy;w_i).\end{aligned}$$

Hence $M_t^nK_\phi(t)$ is a $\mu^{\otimes n}$-martingale.

Finally for each $A\in\mathscr{F}_T^n$ we set

(2.51)
$$\tilde{P}^n(A)=\int_A M_T^n(\mathbf{w})\,\mu^{\otimes n}(dw).$$

Then $\tilde{P}^n$ is well-defined as a probability measure on $(W_T^n,\mathscr{F}_T^n)$. Moreover it is easily seen that for any $\phi\in\mathbb{B}(S^{\otimes n})$

(2.52)
$$K_\phi(t)\ \textit{is a}\ \tilde{P}^n\textit{-martingale}$$

since $M_t^nK_\phi(t)$ is a $\mu^{\otimes n}$-martingale. Therefore $\tilde{P}^n=P^n$, because of the uniqueness of the martingale problem (2.52) with $\mathscr{L}(\mathbf{w}(0);\tilde{P}^n)=u^{\otimes n}$, which completes the proof of Lemma 2.5.

Next we discuss a limit of a functional $H_T^n(w_1,\dots,w_n)$. Since $H_T^n(w_1,\dots,w_n)$ can be regarded as a symmetric statistic of $\{w_n\}_{n=1}^\infty$, which is a sequence of $(W_T,\mathscr{F}_T)$-valued independent random variables with common distribution μ, Dynkin-Mandelbaum's theorem of § 1 is applicable.

Lemma 2.6

(2.53)
$$\lim_{n\to\infty}H_T^n(w_1,\dots,w_n)=\tfrac{1}{2}I_2(f)-\tfrac{1}{2}\mathrm{Trace}\,AA^*$$

in the sense of convergence of probability distributions, where

(2.54) $$f(w,w')=a(w,w')+a(w',w)-\int a(w,w'')\,a(w',w'')\,\mu(dw'')$$

and $I_2(f)$ is a multiple Wiener integral associated with (W_T,μ).

Proof. Using Taylor's expansion of $\log x$ we have

(2.55) $$\begin{aligned} H_T^n(w_1,\dots,w_n) &= \sum_{i=1}^n \int_{[0,T]\times S} \left(\frac{1}{n}\sum_{j=1}^n q_s(w_i(s-),w_j(s-);y)-1\right)\tilde N(dt,dy;w_i)\\ &-\tfrac{1}{2}\sum_{i=1}^n \int_{[0,T]\times S} \left(\frac{1}{n}\sum_{j=1}^n q_s(w_i(s-),w_j(s-);y)-1\right)^2 N(dt,dy;w_i)\\ &+\tfrac{1}{3}\sum_{i=1}^n \int_{[0,T]\times S} \left(\frac{1}{n}\sum_{j=1}^n q_s(w_i(s-),w_j(s-);y)-1\right)^3 N(dt,dy;w_i)\\ &-\tfrac{1}{4}\sum_{i=1}^n \int_{[0,T]\times S} R_i(t)\,N(dt,dy;w_i)\\ &= I_1^n+I_2^n+I_3^n+I_4^n, \end{aligned}$$

where

(2.56) $$|R_i(t)|\leqq \text{const.}\left(\frac{1}{n}\sum_{j=1}^n q_s(w_i(s-),w_j(s-),y)-1\right)^4.$$

We first note

(2.57) $$I_1^n=\frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n a(w_j,w_i).$$

Since it is easily seen that

(2.58) $$a(w,w')\in L^2(\mu\otimes\mu),\quad \text{and}$$

(2.59) $$\int a(w,w')\,\mu(dw')=\int a(w',w)\,\mu(dw)=0 \qquad \mu\text{-a.s.},$$

it follows immediately from Theorem 1.1

(2.60) $$\lim_{n\to\infty} I_1^n=\tfrac{1}{2}I_2(a+a^*).$$

where a^* is defined by $a^*(w,w')=a(w',w)$.

Let

(2.61) $$\begin{aligned} &b(w;w',w'')\\ &=\int_{[0,T]\times S}(q_s(w(s-),w'(s-);y)-1)(q_s(w(s-),w''(s-);y)-1)\,N(ds,dy;w). \end{aligned}$$

Then

$$I_2^n = -\frac{1}{2n^2}\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n b(w_i;w_j,w_k) \tag{2.62}$$

$$= -\frac{1}{2n^2}\sum_{\substack{(i,j,k)\\ \text{distinct}}}(b(w_i;w_j,w_k)-b[\mu;w_j,w_k])$$

$$-\frac{n-2}{2n^2}\sum_{j\neq k} b[\mu;w_j,w_k]-\frac{1}{2n^2}\sum_{i\neq j} b(w_i;w_j,w_j)$$

$$-\frac{1}{n^2}\sum_{i\neq j} b(w_i;w_i,w_j)-\frac{1}{2n^2}\sum_{i=1}^n b(w_i;w_i,w_i)$$

$$= I_{2,1}^n+I_{2,2}^n+I_{2,3}^n+I_{2,4}^n+I_{2,5}^n,$$

where $b[\mu;w',w'']=\int b(w;w',w'')\mu(dw)$. Note that

$$b(w;w',w'')=b(w;w'',w') \quad \textit{and} \quad \int b(w;w',w'')\,d\mu(w'')=0. \tag{2.63}$$

Also it follows from Lemma 2.4

$$b[\mu;w',w'']=\int a(w',w)\,a(w'',w)\,\mu(dw). \tag{2.64}$$

Accordingly, by virtue of Theorem 1.1 we see $I_{2,1}^n I_{2,4}^n$ and $I_{2,5}^n$ vanish as $n\to\infty$, and furthermore we get by (2.64)

$$\lim_{n\to\infty} I_{2,2}^n = -\tfrac{1}{2}I_2(b[\mu;\cdot,\cdot]) \tag{2.65}$$

$$\lim_{n\to\infty} I_{2,3}^n = -\tfrac{1}{2}\int_{W_T} b[\mu;w',w']\,\mu(dw') \tag{2.66}$$

$$= -\tfrac{1}{2}\int_{W_T\times W_T} a(w,w')^2\,\mu(dw)\,\mu(dw')$$

$$= -\tfrac{1}{2}\operatorname{Trace} AA^*.$$

Consequently we obtain

$$\lim_{n\to\infty} I_2^n = \tfrac{1}{2}I_2-\tfrac{1}{2}b[\mu;\cdot,\cdot]-\tfrac{1}{2}\operatorname{Trace} AA^*. \tag{2.67}$$

As for I_3^n and I_4^n, by making use of Theorem 1.1 repeatedly we can show that they both vanish as $n\to\infty$, and the proof of Lemma 2.6 is completed.

Lemma 2.7. Trace $A^n=0$ *for all* $n\geqq 2$.

Proof. By (1.7) it suffices to show that for any $n\geqq 2$

$$\int_{W_T^n} a(w_1,w_2)\,a(w_2,w_3)\dots a(w_n,w_1)\,\mu(dw_1)\dots\mu(dw_n)=0. \tag{2.68}$$

Let

$$a(t;w,w')=\int_{[0,t]\times S}(q_s(w'(s-),w(s-);y)-1)\,\tilde N(ds,dy;w'). \tag{2.69}$$

Since $a(t;w_1,w_2)$, $a(t;w_2,w_3),\ldots,a(t;w_n,w_1)$ are square integrable $\mu^{\otimes n}$ martingales of bounded variation, and any two of them have no common discontinuities with respect to $\mu^{\otimes n}$, Theorem d) of [2] p. 367 implies that $a(t;w_1,w_2)a(t;w_2,w_3)\ldots a(t;w_n,w_1)$ is a $\mu^{\otimes n}$-martingale, which yields (2.68).

Now we are in position to prove Theorem 2.1. Let any $\Phi\in L^2(W_T,\mu)$ be fixed. By Lemma 2.5

$$(2.70)\qquad E^{P^n}[\exp\sqrt{-1}\langle\sqrt{n}(U_n-\mu),\Phi\rangle] = E^{\mu^{\otimes n}}\left[\exp\left\{\sqrt{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\Phi}(w_i)+H_T^n(w_1,\ldots,w_n)\right\}\right].$$

It follows from Lemma 2.6 that $\exp H_T^n(w)$ converges as $n\to\infty$ to $\exp\frac{1}{2}(I_2(f)-\mathrm{Trace}\,AA^*)$ in the law sense. Noting Lemma 2.7 we can apply Lemma 1.2, and we get

$$(2.71)\qquad E[\exp\tfrac{1}{2}(I_2(f)-\mathrm{Trace}\,AA^*)]=1.$$

Hence $\{\exp H_T^n(w)\}$ *is uniformly integrable.* Also, using Theorem 1.1 again

$$\exp\left\{\sqrt{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\Phi}(w_i)+H_T^n(w)\right\}_{n=1}^{\infty}$$

is uniformly integrable and converges as $n\to\infty$ to

$$\exp\{\sqrt{-1}I_1(\tilde{\Phi})+\tfrac{1}{2}I_2(f)-\tfrac{1}{2}\mathrm{Trace}\,AA^*\}$$

in the law sense. Therefore we have by Lemma 1.3

$$(2.72)\qquad \lim_{n\to\infty}E^{\mu^{\otimes n}}\left[\exp\left\{\sqrt{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\Phi}(w_i)+H_T^n(w)\right\}\right] = E[\exp\{\sqrt{-1}I_1(\tilde{\Phi})+\tfrac{1}{2}I_2(f)-\tfrac{1}{2}\mathrm{Trace}\,AA^*\}] = \exp\{-\tfrac{1}{2}\|(I-A)^{-1}\tilde{\Phi}\|^2_{L^2(W_T,\mu)}\},$$

which completes Theorem 2.1.

Next we proceed to prove Theorem 2.2. We can show in the same way as Lemma 2.3 that for any fixed $\zeta\in W_T$ and $\varepsilon>0$ there exists a unique solution of the martingale problem (2.18) and (2.19), which we denote by μ_ε^ζ. Moreover the proof of Lemma 2.5 implies that $\mu_\varepsilon^\zeta\in\mathscr{P}(W_T)$ and $\mu\in\mathscr{P}(W_T)$ are mutually absolutely continuous and

$$(2.73)\qquad \frac{d\mu_\varepsilon^\zeta}{d\mu}(w)=\exp H_T^{\varepsilon,\zeta}(w),$$

where

$$(2.74)\qquad H_T^{\varepsilon,\zeta}(w)=\int_{[0,T]\times S}\{\log q_s[w(s-),u_\varepsilon^\zeta(s)+\varepsilon\delta_{\zeta(s)};y]\}\,N(ds,dy;w) - \int_0^T q[w(s),u_\varepsilon^\zeta(s)+\varepsilon\delta_{\zeta(s)}-u(s)]\,ds$$

where $u_\varepsilon^\zeta(s)=\mathscr{L}(w(s);\mu_\varepsilon^\zeta)$, $u(s)=\mathscr{L}(w(s);\mu)$ and $q_s[x,v;y]$ is defined similarly to $q[x,v]$.

Lemma 2.8. *There exists $C>0$ independent of $\zeta\in W$ and $\varepsilon>0$ such that*

$$\|u_\varepsilon^\zeta(t)-u(t)\|_{\mathrm{var}}\leqq C\varepsilon t \quad \textit{for } 0\leqq t\leqq T, \tag{2.75}$$

and

$$|H_T^{\varepsilon,\zeta}(w)|\leqq C\varepsilon(1+N(T,S;w)) \quad \textit{for any } w\in W_T. \tag{2.76}$$

Proof. For any $\phi\in\mathbb{B}(S)$

$$\begin{aligned}
(2.77)\qquad &|\langle u_\varepsilon^\zeta(t)-u(t),\phi\rangle| \\
&=\left|\int_0^t\langle u_\varepsilon^\zeta(s),Q_{u_\varepsilon^\zeta(s)+\varepsilon\delta_{\zeta(s)}}\phi\rangle\,ds-\int_0^t\langle u(s),Q_{u(s)}\phi\rangle\,ds\right| \\
&=\left|\int_0^t\langle u_\varepsilon^\zeta(s)-u(s),Q_{u(s)}\phi\rangle\,ds+\int_0^t\langle u_\varepsilon^\zeta(s),Q_{u_\varepsilon^\zeta(s)+\varepsilon\delta_{\zeta(s)}-u(s)}\phi\rangle\,ds\right| \\
&\leqq\int_0^t\|u_\varepsilon^\zeta(s)-u(s)\|_{\mathrm{var}}\|Q_{u(s)}\phi\|_\infty\,ds+\int_0^t\|Q_{u_\varepsilon^\zeta(s)+\varepsilon\delta_{\zeta(s)}-u(s)}\phi\|_\infty\,ds \\
&\leqq\text{Const.}\,\|\phi\|_\infty\left(\int_0^t\|u_\varepsilon^\zeta(s)-u(s)\|_{\mathrm{var}}\,ds+\varepsilon t\right).
\end{aligned}$$

Thus we see

$$\|u_\varepsilon^\zeta(t)-u(t)\|_{\mathrm{var}}\leqq\text{Const.}\left(\int_0^t\|u_\varepsilon^\zeta(s)-u(s)\|_{\mathrm{var}}+\varepsilon t\right), \tag{2.78}$$

which yields (2.75). (2.76) follows immediately from (2.75), noting that $q_s(x,x';y)$ is bounded from above and from zero and that a trivial identity $q_s[x,u(s);y]=1$ holds.

We give the proof of Theorem 2.2. We have for any $\Phi\in L^2(W_T,\mu)$

$$\begin{aligned}
(2.79)\qquad \Pi^\varepsilon\Phi(\zeta)&=\frac{1}{\varepsilon}\langle\mu_\varepsilon^\zeta-\mu,\Phi\rangle \\
&=\frac{1}{\varepsilon}\langle\mu,(e^{H_T^{\varepsilon,\zeta}}-1)\Phi\rangle \\
&=\frac{1}{\varepsilon}\langle\mu,H_T^{\varepsilon,\zeta}\Phi\rangle+\frac{1}{\varepsilon}\langle\mu,(e^{H_T^{\varepsilon,\zeta}}-1-H_T^{\varepsilon,\zeta})\Phi\rangle.
\end{aligned}$$

Using Taylor's expansion $H_T^{\varepsilon,\zeta}$ is represented as follows:

$$\begin{aligned}
(2.80)\qquad H_T^{\varepsilon,\zeta}(w)=&\int_{[0,T]\times S}(q_s[w(s-),u_\varepsilon^\zeta(s);y]-1)\tilde{N}(ds,dy;w) \\
&+\varepsilon\int_{[0,T]\times S}(q_s(w(s-),\zeta(s);y)-1)\tilde{N}(ds,dy;w)+J_T^{\varepsilon,\zeta}(w)
\end{aligned}$$

$$|J_T^{\varepsilon,\zeta}(w)|\leqq\text{Const.}\int_{[0,T]\times S}(q_s[w(s-),u_\varepsilon^\zeta(s)+\varepsilon\delta_{\zeta(s)};y]-1)^2N(ds,dy;w). \tag{2.81}$$

Since

$$(2.82)\qquad \begin{aligned}&|q_s[w(s-),u_\varepsilon^\zeta(s)+\varepsilon\delta_{\zeta(s)};y]-1|\\&=|q_s[w(s-),u_\varepsilon^\zeta(s)-u(s);y]-\varepsilon q_s(w(s-),\zeta(s);y)|\\&\leqq\|q_s\|_\infty(\|u_\varepsilon^\zeta(s)-u(s)\|_{\mathrm{var}}+\varepsilon)\end{aligned}$$

it follows from Lemma 2.8 that

$$(2.83)\qquad |J_T^{\varepsilon,\zeta}(w)|\leqq\text{Const.}\,\varepsilon^2 N(T,S;w).$$

Furthermore we note by (2.14) and (2.80)

$$(2.84)\qquad H_T^{\varepsilon,\zeta}(w)=\int a(w',w)(\mu_\varepsilon^\zeta-\mu)(dw')+\varepsilon a(\zeta,w)+J_T^{\varepsilon,\zeta}.$$

Hence combining these it is easily checked

$$(2.85)\qquad \lim_{\varepsilon\downarrow 0}\left(\frac{1}{\varepsilon}\langle\mu,H_T^{\varepsilon,\zeta}\Phi\rangle-\Pi_\varepsilon A\Phi(\zeta)-A\Phi(\zeta)\right)=0$$

uniformly in $\zeta\in W_T$. Also, using an inequality $|e^x-1-x|\leqq\frac{x^2}{2}e^{|x|}$ and (2.76), we have

$$(2.86)\qquad \begin{aligned}&\frac{1}{\varepsilon}|\langle\mu,(e^{H_T^{\varepsilon,\zeta}}-1-H_T^{\varepsilon,\zeta})\Phi\rangle|\\&\leqq\frac{1}{2\varepsilon}\langle\mu,e^{|H_T^{\varepsilon,\zeta}|}|H_T^{\varepsilon,\zeta}|^2|\Phi|\rangle\\&\leqq\frac{1}{2\varepsilon}\|\Phi\|_{L^2(W_T,\mu)}\langle\mu,e^{2|H_T^{\varepsilon,\zeta}|}|H_T^{\varepsilon,\zeta}|^4\rangle^{\frac{1}{2}}\\&\leqq\text{Const.}\,\|\Phi\|_{L^2(W_T,\mu)}\varepsilon\langle\mu,e^{CN(T,S;w)}(1+N(T,S;w))^4\rangle^{\frac{1}{2}},\end{aligned}$$

which tends to 0 as $\varepsilon\downarrow 0$, since $\langle\mu,e^{aN(T,S;w)}\rangle<+\infty$ holds for any $a>0$. Consequently we obtain by (2.79), (2.85) and (2.86)

$$(2.87)\qquad \lim_{\varepsilon\downarrow 0}(\Pi_\varepsilon(I-A)\Phi(\zeta)-A\Phi(\zeta))=0$$

uniformly in $\zeta\in W_T$. Noting that $I-A$ is invertible and $(I-A)^{-1}$ is a bounded operator on $L^2(W_T,\mu)$,

$$(2.88)\qquad s\text{-}\lim_{\varepsilon\downarrow 0}\Pi_\varepsilon\Phi=A(I-A)^{-1}\Phi$$

which completes the proof of Theorem 2.2.

§ 3. The Case of McKean's Model of Boltzmann's Equation

Let us consider McKean's 2-velocity model of Maxwellian gases. For each $v\in\mathscr{P}(\{\pm 1\})$ set

$$(3.1)\qquad Q_v\phi(x)=v(-1)(\phi(-x)-\phi(x))\qquad x=\pm 1.$$

Let any $T>0$ be fixed. W_T denotes the set of all right continuous step functions defined on $[0,T]$ taking values in $\{\pm 1\}$. Let any $u\in\mathscr{P}(\{\pm 1\})$ be given. Then by Lemma 2.3 there exists a unique McKean's measure μ on W_T corresponding to $\{Q_v; v\in\mathscr{P}(\{\pm 1\})\}$, and $u(t)=\mathscr{L}(w(t);\mu)$ is a solution of the following Boltzmann's equation:

$$\frac{d}{dt}\langle u(t),\phi\rangle=\langle u(t)\otimes u(t), Q\phi\rangle, \qquad \phi\in\mathbb{B}(\{\pm 1\}), \tag{3.2}$$

where $Q\phi(x,y)=\phi(xy)-\phi(x)$.

To avoid triviality we henceforth assume $0\leq u(+1)<1$. The n-particle system $X^{(n)}(t)=(X_1^n(t),\dots,X_n^n(t))$ is defined as a Markov process on the n-fold product space $\{\pm 1\}^n$ with generator

$$Q^{(n)}\phi(x_1,\dots,x_n)=\frac{1}{n}\sum_{1\leq i<j\leq n}\sum_c\{\phi(\dots,x_i',\dots,x_j',\dots)-\phi(x_1,\dots,x_n)\} \tag{3.3}$$

where $\sum_c$ is the sum with respect to the two types of collisions

$$\begin{pmatrix}x_i'\\x_j'\end{pmatrix}=\begin{pmatrix}x_ix_j\\x_j\end{pmatrix} \quad\text{or}\quad \begin{pmatrix}x_i\\x_ix_j\end{pmatrix}.$$

It is assumed the initial distribution of $X^{(n)}(t)$ is $u^{\otimes n}$.

Noting

$$Q^{(n)}\phi(\mathbf{x})=\sum_{i=1}^n u_i^{(n)}(\mathbf{x})\{\phi(x_1,\dots,-x_i,\dots,x_n)-\phi(\mathbf{x})\} \qquad (\mathbf{x}=(x_1,\dots,x_n)) \tag{3.4}$$

with $u_i^{(n)}(\mathbf{x})=\frac{1}{n}\sum_{\substack{j=1\\j\neq i}}^n i_{\{-1\}}(x_j)$, we see that this model is essentially contained in the framework of the preceding section although the condition (2.4) fails. If we set for each $w\in W_T$

$$N_t(w)=\sum_{s\leq t} I(w(s)\neq w(s-))$$

$$\tilde N_t(w)=N_t(w)-\int_0^t u(s,-1)\,ds$$

$\tilde N_t(w)$ and $\tilde N_t(w)^2-\int_0^t u(s,-1)\,ds$ are μ-martingales as in Lemma 2.4. We note that Lemma 2.5 should be modified in the following way:

Lemma 2.5′. *Let any $T>0$ be fixed, and let P^n be the probability measure on $W_T^n=W_T\times\dots\times W_T$ induced by $\{X^{(n)}(t)\}_{0\leq t\leq T}$. Then P^n is absolutely continuous with respect to $\mu^{\otimes n}$, and*

$$\frac{dP^n}{d\mu^{\otimes n}}(\mathbf{w})=\exp H_T^n(\mathbf{w}) \tag{3.5}$$

where

$$H_T^n(\mathbf{w}) = \sum_{i=1}^{n} \int_0^T \log \frac{u_i^{(n)}(\mathbf{w}(s-))}{u(s,-1)} dN_s(w_i) - \sum_{i=1}^{n} \int_0^T (u_i^{(n)}(\mathbf{w}(s)) - u(s,-1)) ds.$$

We notice that $-\infty \leqq H_T^n(\mathbf{w}) < +\infty$.

For $(w, w') \in W_T \times W_T$ we set

$$a(w,w') = \int_0^T I_{(-1)}(w(s-)) d\tilde{N}_s(w'), \tag{3.6}$$

and

$$f(w,w') = a(w,w') + a(w',w) - \int_W a(w,w'') a(w' w'') \mu(dw''). \tag{3.7}$$

Then we have

Lemma 2.6′

$$\lim_{n\to\infty} \exp H_T^n(\mathbf{w}) = \exp\{\tfrac{1}{2} I_2(f) - \tfrac{1}{2} \mathrm{Trace}\, A A^*\} \tag{3.8}$$

in the sense of convergence of probability distributions where $I_2(f)$ is a multiple Wiener integral associated with (W_T, μ).

Outline of the proof of the lemma. Let

$$\overline{W}_n = \left\{ \mathbf{w} = (w_1, \ldots, w_n);\ \sup_{0\leqq t\leqq T} \left| \frac{u^{(n)}(\mathbf{w}(t))}{u(s,-1)} - 1 \right| < \frac{1}{2} \right\}.$$

Then we claim

$$\lim_{n\to\infty} \mu^{\otimes n}(\overline{W}_n) = 1. \tag{3.9}$$

Note

$$\sum_{i=1}^{n} \left\{ \phi(w_i(t)) - \langle u(t), \phi\rangle - \int_0^t u(s,-1)(\phi(w_i(s)) - \langle u(s), \phi\rangle) ds \right\}$$

is a $\mu^{\otimes n}$-martingale for any $\phi \in \mathbb{B}(\{\pm 1\})$. Making use of the law of large numbers and a maximal inequality for martingales, it is easy to check

$$\sup_{0\leqq t\leqq T} \left| \frac{1}{n} \sum_{i=1}^{n} \phi(w_i(t)) - \langle u(t), \phi\rangle \right|$$

vanishes as $n \to \infty$ in probability w.r.t. $\mu^{\otimes n}$, and (3.9) follows immediately from this. Accordingly we obtain Lemma 2.6′ combining the proof of Lemma 2.6 with (3.9).

Thus we see that Theorem 2.1 and Theorem 2.2 are valid in the case of McKean's 2-velocity model of Maxwellian gases since the remaining parts of the proof are quite same as the preceding section.

§ 4. The Case of Diffusion Processes

Let $b(x,y)$ be an $\boldsymbol{R}^d$-valued bounded measurable function on $\boldsymbol{R}^d \times \boldsymbol{R}^d$. For each $v \in \mathscr{P}(\boldsymbol{R}^d)$ set

$$Q_v \phi(x) = \frac{1}{2} \sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2} \phi(x) + \sum_{i=1}^{d} b_i[x,v] \frac{\partial}{\partial x_i} \phi(x) \tag{4.1}$$

where $b_i[x,v] = \int b_i(x,y)\, v(dy)\ (i=1,\dots,d)$.

Let any $T>0$ be fixed. Denote by W_T the set of all $\boldsymbol{R}^d$-valued continuous functions defined on $[0,T]$, equipped with the usual σ-fields $\mathscr{F}_t$.

Let any $u \in \mathscr{P}(\boldsymbol{R}^d)$ be given. Then we have

Lemma 2.3″. *There exists a McKean measure μ on W_T corresponding to $\{Q_v; v \in \mathscr{P}(\boldsymbol{R}^d)\}$ with $\mathscr{L}(w(0);\mu)=u$. Namely μ is a solution of the following martingale problem:*

$$\phi(w(t)) - \int_0^t Q_{u(s)} \phi(w(s))\, ds \tag{4.2}$$

is a μ-martingale for any bounded C^2-function ϕ, and

$$\mathscr{L}(w(t);\mu) = u(t) \quad (0 \leqq t \leqq T) \quad \text{and} \quad u(0) = u. \tag{4.3}$$

We omit the proof since it can be shown easily along the same line as in Krylov [5] concerning the existence theorem of SDE. But we think it is not easy to prove the uniqueness of the above martingale problem by some direct methods. However, the method of Sznitman [13] for proving the central limit theorem works even if we do not know the uniqueness in advance and we can easily obtain the uniqueness as a consequence of the central limit theorem.

Now we take a solution μ of the above martingale problem. Then

$$B(t,w) = w(t) - w(0) - \int_0^t b[w(s), u(s)]\, ds, \tag{4.4}$$

is a d-dimensional Brownian motion with respect to μ. The corresponding n-particle system $X^{(n)}(t) = (X_1^n(t), \dots, X_n^n(t))$ is an $(\boldsymbol{R}^d)^n$-valued diffusion process having the initial distribution $u^{\otimes n}$ with generator

$$\begin{aligned} Q^{(n)} \phi(x_1,\dots,x_n) = {} & \tfrac{1}{2} \sum_{i=1}^{n} \Delta_{x_i} \phi(x_1,\dots,x_n) \\ & + \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} b(x_i,x_j) \cdot \nabla_{x_i} \phi(x_1,\dots,x_n) \end{aligned} \tag{4.5}$$

where Δ_{x_t} and ∇_{x_t} act on the i-th variable of $\phi(x_1,\dots,x_n)$. Let P_n be the probability measure on $W_T^n = W_T \times \dots \times W_T$ induced by $\{X^{(n)}(t)\}_{0 \leqq t \leqq T}$.

Lemma 2.5″. *P^n and $\mu^{\otimes n}$ are mutually absolutely continuous, and*

$$\frac{dP^n}{d\mu^{\otimes n}}(\mathbf{w}) = \exp H_T^n(\mathbf{w}) \tag{4.6}$$

where

$$H_T^n(\mathbf{w}) = \sum_{i=1}^{n} \int_0^T \left(\frac{1}{n}\sum_{j=1}^{n} b(w_i(s), w_j(s)) - b[w_i(s), u(s)]\right) dB(s, w_i)$$
$$-\frac{1}{2}\sum_{i=1}^{n}\int_0^T \left\|\frac{1}{n}\sum_{j=1}^{n} b(w_i(s), w_j(s)) - b[w_i(s), u(s)]\right\|^2 ds.$$

For $(w, w') \in W_T \times W_T$ let

(4.7) $$a(w, w') = \int_0^T (b(w'(s), w(s)) - b[w'(s), u(s)])\, dB(s, w')$$

and

(4.8) $$f(w, w') = a(w, w') + a(w', w) - \int_{W_T} a(w, w'')\, a(w', w'')\, \mu(dw'').$$

Then we see that Theorem 2.1 also holds in the present case. In particular, for any bounded measurable function Φ on W_T

$$\lim_{n\to\infty} E^{P^n}[\exp\{\sqrt{-1}\langle U_n, \Phi\rangle\}] = \lim_{n\to\infty} E^{\mu^{\otimes n}}\left[\exp\left\{\frac{\sqrt{-1}}{n}\sum_{i=1}^{n}\Phi(w_i) + H_T^n(\mathbf{w})\right\}\right]$$
$$= \exp\{\sqrt{-1}\langle \mu, \Phi\rangle\},$$

and hence $\left\langle P^n, \frac{1}{n}\sum_{i=1}^{n}\Phi(w_i)\right\rangle \to \langle \mu, \Phi\rangle$ as $n\to\infty$. Since P^n is unique, this implies that $\langle \Phi, \mu\rangle$ is uniquely determined by the initial distribution u. Thus μ is the *unique* solution of the martingale problem (4.2) and (4.3).

Furthermore, by the same reason, for any $\zeta \in W_T$ and $\varepsilon > 0$ there exists a unique probability measure μ_ε^ζ on W_T such that

(4.9) $$\phi(w(t)) - \int_0^t Q_{(u_\varepsilon^\zeta(s) + \delta_{\zeta(s)})}\phi(w(s))\, ds \quad \textit{is a } \mu_\varepsilon^\zeta\text{-martingale}$$

and

(4.10) $$\mathscr{L}(w(t); \mu_\varepsilon^\zeta) = u_\varepsilon^\zeta(t) \quad \text{and} \quad u_\varepsilon^\zeta(0) = u.$$

Noting that μ_ε^ζ and μ are mutually absolutely continuous Π_ε of (2.20) is well-defined, and by a similar argument to §2 we see that Theorem 2.2 is also valid for the diffusion case.

References

1. Dawson, D.A.: Critical dynamics and fluctuations for a mean-field model of cooperative behavior. J. Statist. Phys. **31**, 29–85 (1983)
2. Dellacherie, C., Meyer, P.A.: Probabilities et potentiel, Chapitres V a VIII. Herman: Paris 1980
3. Dynkin, E.B., Mandelbaum, A.: Symmetric statistics, Poisson point processes and multiple Wiener integrals. Ann. Statist. **11**, 739–745 (1983)
4. Ikeda, N., Watanabe, S.: Stochastic differential equations and diffusion processes. Amsterdam: North-Holland/Kodansha 1981

5. Krylov, N.V.: On Ito's stochastic differential equations. Theory Probab. Appl. **14**, 330–336 (1969)
6. Kunita, H., Watanabe, S.: On square integrable martingales. Nagoya Math. J. **30**, 209–245 (1967)
7. Kusuoka, S., Tamura, Y.: Gibbs measures for mean field potentials. J. Fac. Sci. Univ. Tokyo **31**, 223–245 (1984)
8. Maruyama, G.: On the transition probability functions of the Markov processes. Nat. Sci. Rep. Ochanomizu Univ. **5**, 10–20 (1954)
9. McKean, H.P.: Propagation of chaos for a class of non-linear parabolic equations. Lecture Series in Differential Equations, 7 Catholic Univ. 41–57 (1967)
10. McKean, H.P.: Fluctuations in the kinetic theory of gases. Commun. Pure Appl. Math., **28**, 435–455 (1975)
11. Parthasarathy, K.R.: Probability measures on metric spaces. New York: Academic Press 1967
12. Simon, B.: Trace ideals and their applications. London Math. Soc. Lecture Note Series, **35**. Cambridge: Univ. Press 1979
13. Sznitman, A.S.: Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated. J. Funct. Anal. **56**, 311–336 (1984)
14. Sznitman, A.S.: A fluctuation result for non-linear diffusions. [Preprint]
15. Tanaka, H.: Limit theorems for certain diffusion processes with interaction. Proceeding of the Taniguchi International Symposium on Stochastic Analysis (K. Itô ed.), pp. 469–488. Tokyo: Kinokuniya 1984
16. Tanaka, H., Hitsuda, M.: Central limit theorem for a simple diffusion model of interacting particles. Hiroshima Math. J. **11**, 415–423 (1981)

Received October 16, 1984

Probab. Th. Rel. Fields 71, 69–83 (1986)

Probability Theory and Related Fields

Propagation of Chaos for Diffusing Particles of Two Types with Singular Mean Field Interaction

Massao Nagasawa and Hiroshi Tanaka

Institut für Angewandte Mathematik der Universität Zürich, Rämistrasse 74, CH-8001 Zürich, Switzerland

Department of Mathematics, Faculty of Science and Technology, Keio University, Yokohama, 223 Japan

Introduction

In a system of diffusion processes with a pair interaction $b(x, y)$ described by

$$X_i(t)=B_i(t)+\frac{1}{n}\sum_{j=1}^{n}\int_0^t ds\, b(X_i(s), X_j(s)), \qquad i=1,2,\dots,n,$$

where $\{B_i(t)\colon i=1,2,\dots,n\}$ is a family of mutually independent Brownian motions, if the number n of particles tends to infinity, then the distribution of $(X_1, X_2, \dots, X_m)$, for fixed $m \leqq n$, converges to that of a collection of independent copies of a (non-linear) diffusion process $X(\cdot)$ with the so called mean field drift

$$b[x, u(t)]=\int b(x, y)\, u(t, dy)$$

induced by the distribution $u(t,\cdot)$ of $X(t)$ itself. Limit theorems of this kind were called "propagation of chaos" by McKean [4]. One can regard it as a law of large numbers: $u_n=\frac{1}{n}\sum_{i=1}^{n}\delta_{X_i}\to u$. The propagation of chaos for a system of diffusion processes was proved by McKean [5] for Lipschitz continuous interaction $b(x, y)$. Besides this Tanaka [13] discussed an associated fluctuation problem (central limit theorem) modifying a method of Braun and Hepp [1] for Vlasov equation. The same problem has been treated also by Kusuoka and Tamura [2] as a limit theorem of Gibbs measures for mean field potentials, and by Sznitman [11] with the help of (Cameron-Martin) Maruyama formula [3], which enabled him to handle the case of bounded measurable interaction $b(x, y)$ (cf. also Shiga and Tanaka [10]).

The present paper is concerned with the propagation of chaos for a system $\{(X_i, Y_i)\colon i=1,2,\dots,n\}$ of particles of two types with a singular pair-interaction, which arises in a statistical model for segregation of interacting random motion (cf. Nagasawa [6, 7], Nagasawa-Yasue [8] and Nagasawa-Tanaka [9]). It can be described by a system of stochastic differential equations:

$$(1)\quad \begin{aligned} X_i(t) &= X_i(0)+B_i(t)+\int_0^t ds\left\{\frac{1}{n}\sum_{l=1}^{n} f(X_i(s)-Y_l(s))+v(X_i(s))\right\}+\Phi_i(t), \\ &\quad i=1,2,\ldots,n, \\ Y_j(t) &= Y_j(0)+B'_j(t)+\int_0^t ds\left\{-\frac{1}{n}\sum_{l=1}^{n} f(X_l(s)-Y_j(s))+v(Y_j(s))\right\}-\Psi_j(t), \\ &\quad j=1,2,\ldots,n, \end{aligned}$$

where $\{B_i(t), B'_j(t);\, i,j=1,2,\ldots,n\}$ is a family of mutually independent Brownian motions, $X_i(0)$ (resp. $-Y_j(0)$) has a common distribution $\nu^{(+)}$ (resp. $\nu^{(-)}$) whose support is in $[0,\infty)$, $f(x)$ is a non-increasing continuous function on $(0,\infty)$ which may diverge at the origin, an example of which is

$$f(x)=\frac{\alpha}{x^4}, \qquad \alpha>0,$$

$v(x)$ is an odd function which is continuous and nonincreasing in $\mathbb{R}-\{0\}$, and $\Phi_i(t)$ (resp. $\Psi_j(t)$) is a non-decreasing continuous function which makes the origin a reflecting boundary for $X_i(t)$ (resp. $Y_j(t)$). That is, $X_i(t)$ (resp. $Y_j(t)$) is a reflecting diffusion process on $[0,\infty)$ (resp. $(-\infty,0]$) repelled by $Y_l(t)$ with the interaction $\frac{1}{n}f(X_i-Y_l)$ $\left(\text{resp. } -\frac{1}{n}f(X_l-Y_j)\right)$ under the influence of an environment potential $v(X_i)$ (resp. $v(Y_j)$).

The propagation of chaos implies that (X_i, Y_j) converges in law, as $n\to\infty$, to (X, Y) which is a solution of a system of (non-linear) stochastic differential equations

$$(2)\quad \begin{aligned} X(t) &= X(0)+B(t)+\int_0^t ds\{\int_{(-\infty,0]} f(X(s)-y)\,u_Y(s,dy)+v(X(s))\}+\Phi(t), \\ Y(t) &= Y(0)+B'(t)+\int_0^t ds\{-\int_{[0,\infty)} f(x-Y(s))\,u_X(s,dx)+v(Y(s))\}-\Psi(t), \end{aligned}$$

where $u_X(s,\cdot)$ and $u_Y(s,\cdot)$ stand for the distributions of $X(s)$ and $Y(s)$, respectively, $B(t)$ and $B'(t)$ are Brownian motions, and $X(0)$ and $Y(0)$ are distributed by $\nu^{(+)}$ and $\nu^{(-)}$, respectively.

Because of singularity of $f(x)$ (and possibly of $v(x)$) at the origin, it seems that the methods of [2, 5, 11] and [13] can not be immediately applied to the present case. However, thanks to the monotonicity of $f(x)$ and $v(x)$, one can find a rather direct way of proving the propagation of chaos based on the construction of unique solutions of Eqs. (1) and (2) by iteration (Theorem 1 in §1). In Sect. 2, preparing several lemmas, we will prove a limit theorem (Theorem 2). The propagation of chaos, which is a direct corollary of the limit theorem, will be discussed in Sect. 3.

In the original model (cf. [8, 9]) the interaction $f(x)$ (e.g. $f(x)=1/x^4$) diverges at the origin so rapidly that the Y-particles lie always left to the X-particles even without the additional terms Φ_i and $-\Psi_j$, or in other words Y-

particles are separated from the X-particles by a moving boundary. The terms Φ_i and $-\Psi_j$ are added in (1) artificially in order to avoid mathematical difficulties of handling the moving boundary. The propagation of chaos for the unmodified model (i.e. without Φ_i and $-\Psi_j$) is an open problem.

1. Existence and Uniqueness of Solutions

In this section we will prove the existence and uniqueness of solutions of a system of equations

$$(3)\qquad \begin{aligned} \xi(t,w) &= w(t) + \int_0^t ds \int_W b(\xi(s,w), \eta(s,\tilde{z}))\,\mu^{(-)}(d\tilde{z}) + \phi(t,w),\\ \eta(t,z) &= z(t) + \int_0^t ds \int_W b\eta(s,z), \xi(s,\tilde{w}))\,\mu^{(+)}(d\tilde{w}) + \psi(t,z), \end{aligned}$$

where $b(x,y)$ is the sum

$$(4)\qquad b(x,y) = f(x+y) + v(x),$$

of f and v with the properties

(5) $f(x)$ is a non-increasing continuous function on $(0,\infty)$ which may diverge at the origin, and $v(x)$ is an odd function which is nonincreasing and continuous in $\mathbb{R}-\{0\}$.

As will be seen in §3, solutions of the system of Eqs. (3) will provide those for the Eqs. (1) and (2) at the same time.

In Eq. (3), w and z are elements of $W = C([0,\infty)\to\mathbb{R} \cap \{w\colon w(0)\geqq 0\}$, $\mu^{(+)}$ and $\mu^{(-)}$ are probability measures on W, ξ and η are elements of $W^+ = C([0,\infty)\to\mathbb{R}^+)$, $\phi\in\Phi_\xi = C([0,\infty)\to\mathbb{R}^+)\cap\{\phi\colon$ nondecreasing, $\phi(0)=0$ and constant on each connected component of $\{t>0;\, \xi(t)>0\}\}$, and $\psi\in\Phi_\eta$ (cf. [9]).

A dominating process $\zeta(t,w)$ is defined to be the solution of a Skorokhod equation on $[1,\infty)$:

$$(6)\qquad \zeta(t,w) = w(0)\vee 1 + (w(t)-w(0)) + ct + \phi(t,w),$$

where $c\geqq f(1)+v(1)$. The solution (ζ,ϕ) can be given explicitly (cf. §1 of [9]). Because $c\geqq f(x+y)+v(x)$ for $x\geqq 1$ and $y\geqq 0$, Lemma 5 of [9] implies that for any (ξ,η) satisfying (3)

$$(8)\qquad 0\leqq \xi(t,w)\leqq \zeta(t,w),\qquad 0\leqq \eta(t,z)\leqq \zeta(t,z).$$

We impose, in addition, integrability conditions on $\mu^{(+)}$, $\mu^{(-)}$, f and v:

$$(9)\qquad \int_W |w(t)|\,\mu(dw) < \infty,\qquad \mu = \mu^{(+)}, \mu^{(-)};$$

$$(10)\qquad \int_W \int_W f^-(\zeta(s,w)+\zeta(s,z))\,\mu^{(+)}(dw)\,\mu^{(-)}(dz),$$

and

$$\int_W v^-(\zeta(s,w))\,\mu(dw), \qquad \mu=\mu^{(+)},\mu^{(-)}, \tag{11}$$

are bounded on each finite time interval, where $f^- = (-f)\vee 0$ and $v^- = (-v)\vee 0$.

Theorem 1.[1] *Under the conditions* (5), (9), (10) *and* (11) *on* f, v, $\mu^{(+)}$ *and* $\mu^{(-)}$, *there exists a unique solution* $\{(\xi(t,w),\phi(t,w)),(\eta(t,z),\psi(t,z))\}$ *of the system of Eq.* (3), *where* $\xi\in W^+$, $\phi\in\Phi_\xi$, $\eta\in W^+$ *and* $\psi\in\Phi_\eta$.

Proof. Denoting $b^- = (-b)\vee 0$ and

$$\begin{aligned} a^{(+)}(s,x) &= -\int_W b^-(x,\zeta(s,\tilde z))\,\mu^{(-)}(d\tilde z),\\ a^{(-)}(s,x) &= -\int_W b^-(x,\zeta(s,\tilde w))\,\mu^{(+)}(d\tilde w), \end{aligned} \tag{12}$$

we define $(\xi^{(k)}(t,w),\eta^{(k)}(t,z))$, $k\geqq 0$, to be the solutions of Skorokhod equations on $[0,\infty)$

$$\begin{aligned} \xi^{(0)}(t,w) &= w(t)+\int_0^t ds\,a^{(+)}(s,\xi^{(0)}(s,w))+\phi^{(0)}(t,w),\\ \eta^{(0)}(t,z) &= z(t)+\int_0^t ds\,a^{(-)}(s,\eta^{(0)}(s,z))+\psi^{(0)}(t,z), \end{aligned} \tag{13}$$

$$\begin{aligned} \xi^{(k)}(t,w) &= w(t)+\int_0^t ds\int_W b(\xi^{(k)}(s,w),\eta^{(k-1)}(s,\tilde z))\,\mu^{(-)}(d\tilde z)\\ &\quad +\phi^{(k)}(t,w),\\ \eta^{(k)}(t,z) &= z(t)+\int_0^t ds\int_W b(\eta^{(k)}(s,z),\xi^{(k-1)}(s,\tilde w))\,\mu^{(+)}(d\tilde w)\\ &\quad +\psi^{(k)}(t,z). \end{aligned} \tag{14}$$

Theorem 1 of [9] implies the existence of unique solutions $(\xi^{(k)},\phi^{(k)})$ and $(\eta^{(k)},\psi^{(k)})$ which satisfy

$$0\leqq \xi^{(0)}(t,w)\leqq \xi^{(k)}(t,w)\leqq \zeta(t,w), \qquad \text{for all } k\geqq 0, \tag{15}$$

$$\begin{gathered} 0\leqq \xi^{(0)}\leqq \xi^{(2)}\leqq \xi^{(4)}\leqq \ldots \leqq \xi^{(5)}\leqq \xi^{(3)}\leqq \xi^{(1)}\leqq \zeta,\\ \phi^{(1)}\leqq \phi^{(3)}\leqq \phi^{(5)}\leqq \ldots \leqq \phi^{(4)}\leqq \phi^{(2)}\leqq \phi^{(0)}, \end{gathered} \tag{16}$$

and the same inequalities for $(\eta^{(k)},\psi^{(k)})$. Therefore there exist monotone limits

$$\begin{aligned} \underline{\xi}(t,w) &= \lim_{k\to\infty}\xi^{(2k)}(t,w), &\qquad \underline{\phi}(t,w) &= \lim_{k\to\infty}\phi^{(2k)}(t,w),\\ \bar{\xi}(t,w) &= \lim_{k\to\infty}\xi^{(2k+1)}(t,w), &\qquad \bar{\phi}(t,w) &= \lim_{k\to\infty}\phi^{(2k+1)}(t,w), \end{aligned}$$

[1] This is a reformulation of Theorem 2 of [9]

and correspondingly $\underline{\eta}$, $\underline{\psi}$, $\bar{\eta}$ and $\bar{\psi}$. They satisfy

$$\begin{aligned}\underline{\xi}(t,w)&\leqq\bar{\xi}(t,w), & \underline{\phi}(t,w)&\geqq\bar{\phi}(t,w),\\ \underline{\eta}(t,z)&\leqq\bar{\eta}(t,z), & \underline{\psi}(t,z)&\geqq\bar{\psi}(t,z),\end{aligned}\tag{17}$$

$$\begin{aligned}0&\leqq\mu^{(+)}[\underline{\xi}(t,\cdot)]\leqq\mu^{(+)}[\bar{\xi}(t,\cdot)]\leqq\mu^{(+)}[\zeta(t,\cdot)]<\infty,\\ 0&\leqq\mu^{(-)}[\underline{\eta}(t,\cdot)]\leqq\mu^{(-)}[\bar{\eta}(t,\cdot)]\leqq\mu^{(-)}[\zeta(t,\cdot)]<\infty,\end{aligned}\tag{18}$$

$$\begin{aligned}\underline{\xi}(t,w)&=w(t)+\int_0^t ds\int_W b(\bar{\xi}(s,w),\underline{\eta}(s,\tilde{z}))\,\mu^{(-)}(d\tilde{z})+\bar{\phi}(t,w),\\ \bar{\xi}(t,w)&=w(t)+\int_0^t ds\int_W b(\underline{\xi}(s,w),\bar{\eta}(s,\tilde{z}))\,\mu^{(-)}(d\tilde{z})+\underline{\phi}(t,w),\end{aligned}\tag{19}$$

$$\begin{aligned}\bar{\eta}(t,z)&=z(t)+\int_0^t ds\int_W b(\bar{\eta}(s,z),\underline{\xi}(s,\tilde{w}))\,\mu^{(+)}(d\tilde{w})+\bar{\psi}(t,z),\\ \underline{\eta}(t,z)&=z(t)+\int_0^t ds\int_W b(\underline{\eta}(s,z),\bar{\xi}(s,\tilde{w}))\,\mu^{(+)}(d\tilde{w})+\underline{\psi}(t,z).\end{aligned}\tag{20}$$

To get (19) and (20) from (14) notice, for example, $b(\xi^{(2k+1)},\underline{\eta})\leqq b(\xi^{(2k+1)},\eta^{(2k)})\leqq b(\bar{\xi},\eta^{(2k)})$, where the first term (resp. the third) is monotone nondecreasing (resp. nonincreasing).

Then, Lemma 1 below completes the proof of the existence of solutions.

Lemma 1. *If* (17–19) *and* (20) *are fulfilled, then* $\bar{\xi}=\underline{\xi}$, $\bar{\phi}=\underline{\phi}$, $\bar{\eta}=\underline{\eta}$ *and* $\bar{\psi}=\underline{\psi}$.

Proof. Because of the assumptions (5), (10) and (11) it holds that

$$\begin{aligned}&\int_0^t ds\int_W\int_W f^-(\bar{\xi}(s,w)+\underline{\eta}(s,z))\,\mu^{(-)}(dz)\,\mu^{(+)}(dw)\\ &\qquad\leqq\int_0^t ds\int_W\int_W f^-(\zeta(s,w)+\zeta(s,z))\,\mu^{(-)}(dz)\,\mu^{(+)}(dw)<\infty,\\ &\int_0^t ds\int_W v^-(\bar{\xi}(s,w))\,\mu^{(+)}(dw)\leqq\int_0^t ds\int_W v^-(\zeta(s,w))\,\mu^{(+)}(dw)<\infty\end{aligned}$$

and corresponding inequalities for $\bar{\eta}$. Therefore, each term of the right hand side of (19) (resp. (20)) is $\mu^{(+)}$ (resp. $\mu^{(-)}$) integrable because (9) is assumed, and

$$\begin{aligned}0&\leqq\mu^{(+)}[\bar{\xi}(t,\cdot)-\underline{\xi}(t,\cdot)]+\mu^{(-)}[\bar{\eta}(t,\cdot)-\underline{\eta}(t,\cdot)]\\ &\leqq\int_0^t ds\int_W\{v(\bar{\xi}(s,w))-v(\underline{\xi}(s,w))\}\,\mu^{(+)}(dw)+\mu^{(+)}[\bar{\phi}(t,\cdot)-\underline{\phi}(t,\cdot)]\\ &\quad+\int_0^t ds\int_W\{v(\bar{\eta}(s,z))-v(\underline{\eta}(s,z))\}\,\mu^{(-)}(dz)+\mu^{(-)}[\bar{\psi}(t,\cdot)-\underline{\psi}(t,\cdot)],\\ &\leqq\mu^{(+)}[\bar{\phi}(t,\cdot)-\underline{\phi}(t,\cdot)]+\mu^{(-)}[\bar{\psi}(t,\cdot)-\underline{\psi}(t,\cdot)]<\infty.\end{aligned}$$

The inequality above combined with $\bar{\phi}-\underline{\phi}\leqq 0$ and $\bar{\psi}-\underline{\psi}\leqq 0$ implies that $\bar{\phi}(t,w)=\underline{\phi}(t,w)$ for $\mu^{(+)}$-a.e. $w\in W$ (resp. $\bar{\psi}(t,z)=\underline{\psi}(t,z)$ for $\mu^{(-)}$-a.e. $z\in W$) and hence $\bar{\xi}(t,w)=\underline{\xi}(t,w)$ for $\mu^{(+)}$-a.e. $w\in W$ (resp. $\bar{\eta}(t,z)=\underline{\eta}(t,z)$ for $\mu^{(-)}$-a.e. $z\in W$). To show the equalities for all $w, z\in W$, denoting

$$c(s,x)=\int_W b(x,\bar{\eta}(s,z))\,\mu^{(-)}(dz)$$

we consider a Skorokhod equation on $[0,\infty)$

$$\xi(t)=w(t)+\int_0^t ds\,c(s,\xi(s))+\phi(t).$$

Since the solution of the equation is unique by Theorem 1 of [9], $\xi(t,w)=\underline{\xi}(t,w)=\bar{\xi}(t,w)$ and $\phi(t,w)=\underline{\phi}(t,w)=\bar{\phi}(t,w)$ for all $w\in W$. The same argument applies to showing $\eta(t,z)=\underline{\eta}(t,z)=\bar{\eta}(t,z)$ for all $z\in W$. This completes the proof of Lemma 1.

To prove the uniqueness let $(\xi_1,\phi_1,\eta_1,\psi_1)$ and $(\xi_2,\phi_2,\eta_2,\psi_2)$ be solutions of Eq. (3) and define sequences $\{\xi^{(k)}\}$ and $\{\eta^{(k)}\}$ by

$$\xi^{(0)}=\xi_1\wedge\xi_2 \qquad \eta^{(0)}=\eta_1\wedge\eta_2,$$

and

$$\begin{aligned}\xi^{(k)}(t,w)&=w(t)+\int_0^t ds\int_W b(\xi^{(k)}(s,w),\eta^{(k-1)}(s,\tilde{z}))\,\mu^{(-)}(d\tilde{z})\\&\quad+\phi^{(k)}(t,w),\\ \eta^{(k)}(t,z)&=z(t)+\int_0^t ds\int_W b(\eta^{(k)}(s,z),\xi^{(k-1)}(s,\tilde{w}))\,\mu^{(+)}(d\tilde{w})\\&\quad+\psi^{(k)}(t,z),\end{aligned}$$

for $k\geqq 1$. Then it is easy to see that

$$\begin{aligned}&\xi^{(2k)}\leqq\xi_1,\xi_2\leqq\xi^{(2k+1)}, &&\text{for } k\geqq 0,\\ &\xi^{(0)}\geqq\xi^{(2)}\geqq\xi^{(4)}\geqq\dots, &&\phi^{(2)}\leqq\phi^{(4)}\leqq\phi^{(6)}\leqq\dots,\\ &\xi^{(1)}\leqq\xi^{(3)}\leqq\xi^{(5)}\leqq\dots, &&\phi^{(1)}\geqq\phi^{(3)}\geqq\phi^{(5)}\geqq\dots,\end{aligned}$$

and the same inequalities for $(\eta^{(k)},\psi^{(k)})$, which imply the existence of monotone limits

$$\begin{aligned}\bar{\xi}(t,w)&=\lim_{k\to\infty}\xi^{(2k+1)}(t,w), &\bar{\phi}(t,w)&=\lim_{k\to\infty}\phi^{(2k+1)}(t,w),\\ \underline{\xi}(t,w)&=\lim_{k\to\infty}\xi^{(2k)}(t,w), &\underline{\phi}(t,w)&=\lim_{k\to\infty}\phi^{(2k)}(t,w),\end{aligned}$$

and correspondingly $(\bar{\eta},\bar{\psi},\underline{\eta},\underline{\psi})$. It is clear that they satisfy (17–19) and (20). Therefore, Lemma 1 implies $\bar{\xi}=\underline{\xi}$, $\bar{\phi}=\underline{\phi}$, $\bar{\eta}=\underline{\eta}$, $\bar{\psi}=\underline{\psi}$, and hence $\xi_1=\xi_2$, $\phi_1=\phi_2$, $\eta_1=\eta_2$ and $\psi_1=\psi_2$, completing the proof of Theorem 1.

2. A Limit Theorem

Given a function $b(t, x)$ defined on $[0, \infty) \times (0, \infty)$ with the property

(21) $b(t, x)$ is continuous in $(t, x) \in [0, \infty) \times (0, \infty)$ and nonincreasing in x for each $t \geqq 0$,

we consider a Skorokhod equation on $[0, \infty)$

$$\xi(t, w) = w(t) + \int_0^t ds\, b(s, \xi(s, w)) + \phi(t, w). \tag{22}$$

Theorem 1 of [9] guarantees the existence of a unique solution $(\xi(t, w), \phi(t, w))$ for $w \in W$, where $\xi \in W^+$ and $\phi \in \Phi_\xi$. We will denote it as $(\xi(t), \phi(t))$ without w, if there will be no confusion.

Lemma 2. *Let $b(t, x)$ satisfy the condition* (21) *and let (ξ^a, ϕ^a), $a > 0$, be the solution of a Skorokhod equation on $[a, \infty)$:*

$$\xi^a(t) = a + w(t) + \int_0^t ds\, b(s, \xi^a(s)) + \phi^a(t). \tag{23}$$

Then

$$\xi^a(t) \searrow \xi(t), \quad as \quad a \searrow 0.$$

Proof. See Lemma 5 and what follows in Sect. 3 of [9].

Lemma 3. *Under the assumption* (21) *define for $a > 0$*

$$b_a(t, x) = b(t, x \vee a), \quad for\ (t, x) \in [0, \infty) \times (0, \infty). \tag{24}$$

Let (ξ_a, ϕ_a) be the solution of (22) *with $b_a(t, x)$ in place of $b(t, x)$. Then*

$$\xi_\alpha(t) \nearrow \xi(t), \quad as \quad a \searrow 0.$$

Proof. Theorem 1 of [9] implies that $\xi_a(t) \nearrow$ and $\phi_a(t) \searrow$, as $a \searrow 0$. Therefore

$$\int_0^t ds\, b_a(s, \xi_a(s)) = \xi_a(t) - w(t) - \phi_a(t) \nearrow, \quad \text{as} \quad a \searrow 0,$$

and hence the limit

$$\lim_{a \searrow 0} \int_0^t ds\, b_a(s, \xi_a(s))$$

exists. Denote

$$\tilde{\xi}(t) = \lim_{a \searrow 0} \xi_a(t).$$

Since $b_a(s, \xi_a(s))\, 1_{(\varepsilon, \infty)}(\xi_a(s))$ decreases to $b(s, \tilde{\xi}(s))\, 1_{(\varepsilon, \infty)}(\tilde{\xi}(s))$, as $a \searrow 0$, for $\varepsilon > 0$, we have

$$\lim_{a\searrow 0}\int_0^t ds\, b_a(s,\xi_a(s))$$
$$=\int_0^t ds\, b(s,\tilde\xi(s))\,1_{(\varepsilon,\infty)}(\tilde\xi(s))+\lim_{a\searrow 0}\int_0^t ds\, b_a(s,\xi_a(s))\,1_{[0,\varepsilon]}(\xi_a(s))$$
$$=\int_0^t ds\, b(s,\tilde\xi(s))\,1_{(0,\infty)}(\tilde\xi(s))+\lim_{\varepsilon\searrow 0}\lim_{a\searrow 0}\int_0^t ds\, b_a(s,\xi_a(s))\,1_{[0,\varepsilon]}(\xi_a(s)).$$

Therefore $\tilde\xi(t)$ satisfies

$$\tilde\xi(t)=w(t)-ct+\int_0^t ds\,\{b(s,\tilde\xi(s))+c\}+\tilde\phi(t), \tag{25}$$

where a constant c is chosen to be $b(t,1)+c>0$, for $\forall t\in[0,T]$, and $\tilde\phi$ is defined by

$$\tilde\phi(t)=\lim_{\varepsilon\searrow 0}\lim_{a\searrow 0}\int_0^t ds\,\{b_a(s,\xi_a(s))+c\}\,1_{[0,\varepsilon]}(\xi_a(s))+\lim_{a\searrow 0}\phi_a(t).$$

It is easy to see that $\tilde\xi(t)$ is continuous in t (cf. §3 of [9]), and hence $\tilde\phi(t)$ is also continuous in t by (25). Moreover, $\tilde\phi(t)$ is monotone nondecreasing and stays constant on each connected open interval in which $\tilde\xi(t)>0$. Because of the uniqueness of solutions for a Skorokhod equation (22) which is the same as (25), the $(\tilde\xi,\tilde\phi)$ must coincide with (ξ,ϕ), completing the proof.

We consider W and W^+ as metric spaces with the metric of the uniform convergence on each finite time interval.

Lemma 4. *Let $b(t,x)$ be bounded and Lipschitz continuous: $|b(t,x)-b(t,y)|\leq c(T)\cdot|x-y|$, for $\forall t\in[0,T]$, and $\forall x,y\in[0,\infty)$. Then the solution $\xi(t,w)$ of Eq.* (22) *is continuous in $(t,w)\in[0,T]\times W$.*

Proof. The lemma is an easy consequence of Lemma 2 of [9]. In fact, as is shown there, the solution $\xi(t,w)$ is given as the limit of $\{\xi^{(k)}(t,w)\}$ defined by

$$\begin{aligned}\xi^{(0)}(t,w)&=w(t)+\phi^{(0)}(t,w),\\ \xi^{(k)}(t,w)&=w(t)+\int_0^t ds\, b(s,\xi^{(k-1)}(s,w))+\phi^{(k)}(t,w),\end{aligned} \tag{26}$$

and the convergence is uniform in $(t,w)\in[0,T]\times W$. Since it is easy to see by induction that $\xi^{(k)}(t,w)$ is continuous in $(t,w)\in[0,T]\times W$, so is $\xi(t,w)$, completing the proof.

Lemma 5. *Under the condition* (21) *on $b(t,x)$, the solution $\xi(t,w)$ of Eq.* (22) *is continuous in $(t,w)\in[0,T]\times W$, for any $\forall T>0$.*

Proof. Let K be an arbitrary compact subset of W and denote $M=\sup_{w\in K}\sup_{t\in[0,T]}\zeta(t,w)<\infty$, where ζ is a dominating process which is the solution of a Skorokhod equation on $[1,\infty)$

$$\zeta(t,w)=w(0)\vee 1+(w(t)-w(0))+ct+\psi(t,w)$$

with $c \geqq \sup_{t \in [0, T]} b(t, 1)$. It is to be noted that $b_a(t, x)$ (resp. $b(t, x)$) can be approximated from below (resp. from above) on $[0, M+1]$ (resp. $[a, M+1]$) by Lipschitz continuous ones. Therefore, $\xi_a(t, w)$ (resp. $\xi^a(t, w)$) is the increasing (resp. decreasing) limit of continuous functions on $[0, T] \times K$, and hence lower (resp. upper) semi-continuous in $(t, w) \in [0, T] \times K$. Now, let $a \searrow 0$. Then $\xi_a(t, w)$ (resp. $\xi^a(t, w)$) increases (resp. decreases) to $\xi(t, w)$ by Lemma 2 and 3. Therefore $\xi(t, w)$ is lower and upper semi-continuous, and hence continuous in $(t, w) \in [0, T] \times K$. This completes the proof of Lemma 5.

Lemma 6. *Let $b(t, x)$ and $b_n(t, x)$, $n = 1, 2, 3, \ldots$, satisfy the condition* (21), *and for* $\forall a > 0$ *and* $\forall T > 0$

$$\text{(27)} \qquad \begin{aligned} &\lim_{n \to \infty} b_n(t, x) = b(t, x), \text{ uniformly in } (t, x) \in [0, T] \times \left[a, \frac{1}{a}\right], \\ &\sup_{n \geqq 1} \max_{t \in [0, T]} b_n(t, a) < \infty. \end{aligned}$$

Let $\xi_n(t, w)$ denote the solution of Eq. (22) *with $b_n(t, x)$ in place of $b(t, x)$. Then for any compact subset K of W*

$$\text{(28)} \qquad \lim_{n \to \infty} \xi_n(t, w) = \xi(t, w), \quad \text{uniformly in} \quad (t, w) \in [0, T] \times K.$$

Proof. We first prove the assertion assuming that $b(t, x)$ is bounded above and $b_n(t, x)$ converges uniformly on $[0, T] \times \left[0, \frac{1}{a}\right]$. With the help of Lemma 1 of [9] we have

$$\begin{aligned} |\xi_n(t) - \xi(t)|^2 \leqq\, & 2 \int_0^t ds (\xi_n(s) - \xi(s)) \{b_n(s, \xi_n(s)) - b_n(s, \xi(s))\} \\ & + 2 \int_0^t ds (\xi_n(s) - \xi(s)) \{b_n(s, \xi(s)) - b(s, \xi(s))\}, \end{aligned}$$

where the first integral is nonpositive and the second is, by Schwartz inequality, dominated by

$$\int_0^t ds |\xi_n(s) - \xi(s)|^2 + \int_0^t ds |b_n(s, \xi(s)) - b(s, \xi(s))|^2.$$

Therefore, applying Gronwall's lemma, we have

$$\text{(29)} \qquad |\xi_n(t, w) - \xi(t, w)|^2 \leqq e^T \int_0^T ds |b_n(s, \xi(s, w)) - b(s, \xi(s, w))|^2.$$

Since

$$\sup_{w \in K} \sup_{s \in [0, T]} \xi(s, w) \leqq \sup_{w \in K} \sup_{s \in [0, T]} \zeta(s, w) < \infty,$$

for any compact subset K of W, the integrand of the right hand side of (29) is bounded and converges to zero uniformly in $(s, w) \in [0, T] \times K$, as $n \to \infty$. Therefore (29) implies that $\xi_n(t, w)$ converges to $\xi(t, w)$ uniformly on $[0, T] \times K$.

We apply what we have proved above to $b_a(t,x)$ and $(b_n)_a(t,x)$ defined by (24) (resp. $b(t,x)$ and $b_n(t,x)$ for $x\geqq a$) and obtain, using the notations in Lemma 2 and 3,

$$(30)\qquad (\xi_n)_a(t,w)\to\xi_a(t,w),\quad (\xi_n)^a(t,w)\to\xi^a(t,w),\quad \text{as}\quad n\to\infty,$$

uniformly on $[0,T]\times K$. On the other hand $\xi_a(t,w)\nearrow\xi(t,w)$ and $\xi^a(t,w)\searrow\xi(t,w)$ as $a\searrow 0$ uniformy on $[0,T]\times K$ by Lemma 2 and Dini's theorem, and hence there exists $a>0$ such that

$$0\leqq\xi^a(t,w)-\xi_a(t,w)<\varepsilon,\quad \text{uniformly in}\quad (t,w)\in[0,T]\times K.$$

Because of (30) it holds that for this $a>0$

$$|(\xi_n)_a(t,w)-\xi_a(t,w)|<\varepsilon,\qquad |(\xi_n)^a(t,w)-\xi^a(t,w)|<\varepsilon,$$

uniformly on $[0,T]\times K$, for sufficiently large n. Therefore we have

$$|\xi(t,w)-\xi_n(t,w)|\leqq 3\varepsilon,\quad \text{uniformly in}\quad (t,w)\in[0,T]\times K,$$

where made is use of the inequality $(\xi_n)_a\leqq\xi_n\leqq(\xi_n)^a$. This completes the proof of Lemma 6.

Theorem 2. *Besides* (5) *assume that f is bounded below.*[2] *Let* $\mu^{(+)}$, $\mu^{(-)}$, $\mu_n^{(+)}$, $\mu_n^{(-)}$, $n=1,2,\ldots$, *be probability measures on W satisfying conditions* (9) *and* (11) *for given v. Let* (ξ,η) *(resp.* (ξ_n,η_n)*) be solution of* (3) *with* $(\mu^{(+)},\mu^{(-)})$ *(resp.* $(\mu_n^{(+)},\mu_n^{(-)})$*). If* $\mu_n^{(+)}$ *(resp.* $\mu_n^{(-)}$*) converges to* $\mu^{(+)}$ *(resp.* $\mu^{(-)}$*) weakly, then for any compact subsets* K_1, K_2 *of W and* $\forall\, T>0$

$$(31)\qquad \lim_{n\to\infty}\xi_n(t,w)=\xi(t,w),\qquad \lim_{n\to\infty}\eta_n(t,z)=\eta(t,z)$$

uniformly in $(t,w,z)\in[0,\mathrm{T}]\times K_1\times K_2$.

Proof. Let $(\xi^{(k)},\eta^{(k)})$ and $(\xi_n^{(k)},\eta_n^{(k)})$, $k\geqq 0$, be defined by (12), (13) and (14) with $(\mu^{(+)},\mu^{(-)})$ and $(\mu_n^{(+)},\mu_n^{(-)})$, respectively. We will prove first of all that $\xi_n^{(k)}(t,w)$ (resp. $\eta_n^{(k)}(t,z)$) converges to $\xi^{(k)}(t,w)$ (resp. $\eta^{(k)}(t,z)$), as $n\to\infty$, uniformly on $[0,T]\times K$, where K is a compact subset of W. We show it by induction in k. Since $\mu_n^{(+)}$ (resp. $\mu_n^{(-)}$) converges weakly to $\mu^{(+)}$ (resp. $\mu^{(-)}$) by the assumption, $(a^{(+)}(s,x), a^{(-)}(s,x))$ and $(a_n^{(+)}(s,x), a_n^{(-)}(s,x))$ defined by (12) with $(\mu^{(+)},\mu^{(-)})$ and $(\mu_n^{(+)},\mu_n^{(-)})$, respectively, satisfy the conditions of Lemma 6. Therefore, $\xi_n^{(0)}(t,w)$ (resp. $\eta_n^{(0)}(t,z)$) converges to $\xi^{(0)}(t,w)$ (resp. $\eta^{(0)}(t,z)$) uniformly on $[0,T]\times K$.

Assume that $\xi_n^{(k)}(t,w)$ (resp. $\eta_n^{(k)}(t,z)$) converges to $\xi^{(k)}(t,w)$ (resp. $\eta^{(k)}(t,z)$), as $n\to\infty$, uniformly on $[0,T]\times K$. We will only prove that $\xi_n^{(k+1)}(t,w)$ converges to $\xi^{(k+1)}(t,w)$ uniformly in $(t,w)\in[0,T]\times K$, since the same argument can be applied to $\eta_n^{(k+1)}(t,z)$ and $\eta^{(k+1)}(t,z)$. Define

$$b^{(k+1)}(s,x)=\int_W f(x+\eta^{(k)}(s,\tilde z))\,\mu^{(-)}(d\tilde z)+v(x),$$

$$b_n^{(k+1)}(s,x)=\int_W f(x+\eta_n^{(k)}(s,\tilde z))\,\mu_n^{(-)}(d\tilde z)+v(x).$$

---

[2] Then f satisfies the condition (10) automatically

It is clear that $\xi^{(k+1)}$ and $\xi_n^{(k+1)}$ satisfy Skorokhod equations

$$\xi^{(k+1)}(t, w) = w(t) + \int_0^t ds\, b^{(k+1)}(s, \xi^{(k+1)}(s, w)) + \phi^{(k+1)}(t, w),$$

$$\xi_n^{(k+1)}(t, w) = w(t) + \int_0^t ds\, b_n^{(k+1)}(s, \xi_n^{(k+1)}(s, w)) + \phi_n^{(k+1)}(t, w),$$

respectively. If we prove that $b^{(k+1)}(t, x)$ and $b_n^{(k+1)}(t, x)$ satisfy the conditions of Lemma 6, it implies that $\xi_n^{(k+1)}(t, w)$ converges to $\xi^{(k+1)}(t, w)$ uniformly on $[0, T] \times K$, *as* $n \to \infty$.

The condition (21) is clearly satisfied. The condition (27) can be verified as follows:

$$(32) \qquad \begin{aligned} &|b_n^{(k+1)}(s, x) - b^{(k+1)}(s, x)| \\ &\leqq \int_W |f(x + \eta_n^{(k)}(s, z)) - f(x + \eta^{(k)}(s, z))|\, \mu_n^{(-)}(dz) \\ &\quad + \left|\int_W f(x + \eta^{(k)}(s, z))(\mu_n^{(-)}(dz) - \mu^{(-)}(dz))\right|. \end{aligned}$$

Since $\mu_n^{(-)}$ converges weakly to $\mu^{(-)}$ by the assumption, there exists a compact subset K_0 of W such that

$$\mu_n^{(-)}(K_0^c) \leqq \frac{\varepsilon}{2f(a)}.$$

Keeping this in mind, we estimate

(33) The first term of the right hand side of (32)
$$\leqq \int_{K_0} |f(x + \eta_n^{(k)}(s, z)) - f(x + \eta^{(k)}(s, z))|\, \mu_n^{(-)}(dz) + \varepsilon,$$

where the integral is smaller than ε for sufficiently large n uniformly in $(s, x) \in [0, T] \times \left[a, \frac{1}{a}\right]$, because the integrand converges to zero uniformly in $(s, x, z) \in [0, T] \times \left[a, \frac{1}{a}\right] \times K_0$; and

(34) The second term of the right hand side of (32)
$$\leqq \left|\int_{K_0} f(x + \eta^{(k)}(s, z))(\mu_n^{(-)}(dz) - \mu^{(-)}(dz))\right| + \varepsilon,$$

where $f(x + \eta^{(k)}(s, z))$ is uniformly continuous in $(s, x, z) \in [0, T] \times \left[a, \frac{1}{a}\right] \times K_0$, and the integral is smaller than 3ε uniformly in $(s, x) \in [0, T] \times \left[a, \frac{1}{a}\right]$ for sufficiently large n. In fact,[3] let $\left\{U(s, x) : (s, x) \in [0, T] \times \left[a, \frac{1}{a}\right]\right\}$ be a covering of

[3] Thanks to a referee's remark one can invoke Theorem 6.8 (p. 51) in *Probability measures on metric spaces* by Parthasarathy

$[0, T] \times \left[a, \frac{1}{a}\right]$, where

$$U(s, x) = \{(t, y): \sup_{z \in K_0} |f(x + \eta^{(k)}(s, z)) - f(y + \eta^{(k)}(t, z))| < \varepsilon\}.$$

Then, there exists a finite subcovering $\{U(s_i, x_i): i = 1, 2, \ldots, l\}$. Let n be sufficiently large so that

$$\left|\int_{K_0} f(x_i + \eta^{(k)}(s_i, z))(\mu_n^{(-)}(dz) - \mu^{(-)}(dz))\right| < \varepsilon$$

for $i = 1, 2, \ldots, l$. For any $(s, x) \in [0, T] \times \left[a, \frac{1}{a}\right]$ we can find $U(s_i, x_i)$ such that $(s, x) \in U(s_i, x_i)$, and hence

$$\begin{aligned} &\left|\int_{K_0} f(x + \eta^{(k)}(s, z))(\mu_n^{(-)}(dz) - \mu^{(-)}(dz))\right| \\ &\quad < 2\varepsilon + \left|\int_{K_0} f(x_i + \eta^{(k)}(s_i, z))(\mu_n^{(-)}(dz) - \mu^{(-)}(dz))\right| \\ &\quad < 3\varepsilon, \end{aligned}$$

for sufficiently large n. Thus we have finally

$$\sup_{(s,x) \in [0, T] \times \left[a, \frac{1}{a}\right]} |b_n^{(k+1)}(s, x) - b^{(k+1)}(s, x)| \leqq 6\varepsilon$$

for sufficiently large n. Hence $b^{(k+1)}(s, x)$ and $b_n^{(k+1)}(s, x)$ satisfy the condition (27), the second condition of (27) being trivial.

Now, the proof can be completed as follows: Because of Theorem 1, especially inequality (16), Lemma 5 and Dini's theorem we can find k_0 for any $\varepsilon > 0$ such that

$$0 \leqq \xi^{(2k+1)}(t, w) - \xi^{(2k)}(t, w) \leqq \varepsilon, \quad \text{for } k \geqq k_0, \tag{35}$$

uniformly in $(t, w) \in [0, T] \times K$. On the other hand there exists n_0 such that for $n \geqq n_0$

$$\begin{aligned} &|\xi_n^{(2k)}(t, w) - \xi^{(2k)}(t, w)| \leqq \varepsilon, \\ &|\xi_n^{(2k+1)}(t, w) - \xi^{(2k+1)}(t, w)| \leqq \varepsilon, \end{aligned} \tag{36}$$

uniformly in $(t, w) \in [0, T] \times K$ as is proved above. Therefore, combining (35) and (36) with

$$\begin{aligned} &\xi^{(2k)}(t, w) \leqq \xi(t, w) \leqq \xi^{(2k+1)}(t, w), \\ &\xi_n^{(2k)}(t, w) \leqq \xi_n(t, w) \leqq \xi_n^{(2k+1)}(t, w), \end{aligned} \tag{37}$$

we obtain

$$\sup_{(t, w) \in [0, T] \times K} |\xi(t, w) - \xi_n(t, w)| \leqq 3\varepsilon, \quad \text{for } n \geqq n_0.$$

Thus $\xi_n(t,w)$ converges to $\xi(t,w)$, as $n\to\infty$, uniformly on $[0,T]\times K$ for any $T>0$ and compact subset K of W. This completes the proof of Theorem 2.

3. Propagation of Chaos

We consider a system of Eq. (3) with f and v satisfying the condition (5). Moreover we assume that f is bounded below. Let P_x be the Wiener measure on W starting from $x\in\mathbb{R}$, $\nu^{(+)}$ and $\nu^{(-)}$ be probability measures on $\mathbb{R}^+$ with

(38) $$\int_{\mathbb{R}^+} x\,\nu(dx)<\infty, \qquad \nu=\nu^{(+)},\nu^{(-)},$$

and define probability measures $\mu^{(+)}$ and $\mu^{(-)}$ on W by

(39) $$\mu^{(+)}(\cdot)=\int_{\mathbb{R}^+}\nu^{(+)}(dx)P_x(\cdot), \qquad \mu^{(-)}(\cdot)=\int_{\mathbb{R}^+}\nu^{(-)}(dx)P_x(\cdot).$$

Because of (38), the probability measures $\mu^{(+)}$ and $\mu^{(-)}$ satisfy the condition (9). In addition we assume the condition (11).[4] Then there exists a unique solution $(\xi(t,w),\eta(t,z))$ of (3) with $(\mu^{(+)},\mu^{(-)})$ for all $(w,z)\in W\times W$ by Theorem 1. A stochastic process (X,Y) on $(W\times W,\mu^{(+)}\otimes\mu^{(-)})$ defined by

(40) $$(X(t),Y(t))=(\xi(t),-\eta(t)),$$

coincides clearly with the process described by Eq. (2).

Let (Ω,P) be the infinite product of $(W\times W,\mu^{(+)}\otimes\mu^{(-)})$ i.e.

(41) $$\Omega=W^2\times W^2\times\ldots, \qquad W^2=W\times W,$$
$$P=\mu\otimes\mu\otimes\ldots, \qquad \mu=\mu^{(+)}\otimes\mu^{(-)}.$$

For $\omega=((w_1,z_1),(w_2,z_2),\ldots)\in\Omega$ define $\mathbf{X}(t,\omega)$ by

(42) $$X_i(t,\omega)=\xi(t,w_i),$$
$$Y_i(t,\omega)=-\eta(t,z_i), \qquad i=1,2,3,\ldots,$$

and

(43) $$\mathbf{X}(t)=((X_1(t),Y_1(t)),(X_2(t),Y_2(t)),\ldots).$$

By definition $\{(X_i(t),Y_i(t))\colon i=1,2,3,\ldots\}$ is a family of independent copies of the process described by (2).

Now, define for $\omega=((w_1,z_1),(w_2,z_2),\ldots)\in\Omega$ sequences $\{\mu_n^{(+)}(\omega)\}$ and $\{\mu_n^{(-)}(\omega)\}$ of probability measures on W by

(44) $$\mu_n^{(+)}(\omega,\cdot)=\frac{1}{n}\sum_{i=1}^n\delta_{w_i}(\cdot),$$
$$\mu_n^{(-)}(\omega,\cdot)=\frac{1}{n}\sum_{i=1}^n\delta_{z_i}(\cdot).$$

---

[4] A sufficient condition for (11) is given in Proposition 1 of [9], i.e. if v^- satisfies (i) $v^-(x)\leq c(1+|x|^\alpha)$ for some nonnegative constants c and α, and if $\nu=\nu^{(+)},\nu^{(-)}$ satisfy (ii) $\int_{\mathbb{R}^+}x^\alpha\nu(dx)<\infty$, then the condition (11) is fullfilled

They satisfy trivially the conditions (9) and (11). Therefore, by Theorem 1, there exists for each $(w,z)\in W\times W$ a unique solution $(\xi_n(t,w,\omega),\eta_n(t,z,\omega))$ of (3) with $(\mu_n^{(+)}(\omega),\mu_n^{(-)}(\omega))$. Define, for $\omega=((w_1,z_1),(w_2,z_2),\ldots)\in\Omega$, processes $(X_i^{(n)}, Y_i^{(n)})$, $i=1,2,\ldots,n$, on (Ω,P) by

$$X_i^{(n)}(t,\omega)=\xi_n(t,w_i,\omega),\qquad Y_i^{(n)}(t,\omega)=-\eta_n(t,z_i,\omega), \tag{45}$$

and put

$$\mathbf{X}^{(n)}(t)=((X_1^{(n)}(t),Y_1^{(n)}(t)),\ldots,(X_n^{(n)}(t),Y_n^{(n)}(t))). \tag{46}$$

Clearly $\{(X_i^{(n)}(t),Y_i^{(n)}(t)):\ i=1,2,\ldots,n\}$ is a family of particles of two types described by (1).

Since $\mu_n^{(+)}(\omega)$ (resp. $\mu_n^{(-)}(\omega)$) converges to $\mu^{(+)}$ (resp. $\mu^{(-)}$) weakly for P-a.e. $\omega\in\Omega$ by the strong law of large numbers, Theorem 2 implies that for P-a.e. $\omega\in\Omega$ the sequence $(\xi_n(t,w,\omega),\ \eta_n(t,z,\omega))$ converges to $(\xi(t,w),\eta(t,z))$ uniformly in $(t,w,z)\in[0,T]\times K_1\times K_2$, for $\forall T>0$ and $\forall$ compact $K_1,K_2\subset W$. In other words, the propagation of chaos (of Kac-McKean) holds for a family of diffusing particles of two types with singular interaction described by (1).

Theorem 3. (Propagation of chaos.) *Under the stated conditions*[5] *let* $((X_1^{(n)}(t,\omega),Y_1^{(n)}(t,\omega)),\ldots,(X_n^{(n)}(t,\omega),Y_n^{(n)}(t,\omega)))$ *be the system of* $2n$*-particles of two types defined on* (Ω,P) *by* (45) *and* (46). *Then, for* P-a.e. $\omega\in\Omega$, $((X_1^{(n)}(t,\omega),Y_1^{(n)}(t,\omega)),\ldots,(X_m^{(n)}(t,\omega),\ Y_m^{(n)}(t,\omega)))$, *for fixed* $m\leqq n$, *converges uniformly in* $t\in[0,T]$ *for* $\forall T>0$, *as* $n\to\infty$, *to* $((X_1(t,\omega),Y_1(t,\omega)),\ldots,X_m(t,\omega),Y_m(t,\omega)))$ *which is a family of independent copies of the non-linear diffusion process of two types defined by* (42). *That is, the diffusion process on* $(\mathbb{R}^2)^n$ *described by equation* (1) *converges in law to the infinite direct product of the diffusion process described by* (2) *as n tends to infinity.*

Remark. P_x in (39) is not necessarily to be the Wiener measure, and Theorem 3 remains valid for arbitrary Markov processes on W, since our treatment of the problem is purely pathwise and does not depend on a special choice of probability measures on W.

[5] See the first paragraph of this section

References

1. Braun, W., K. Hepp: The Vlasov dynamics and its fluctuations in $1/N$ limit of interacting classical particles. Commun. Math. Phys. **56**, 101–113 (1977)
2. Kusuoka, S., Y. Tamura: Gibbs measures for mean field potentials. J. Fac. Sci. Univ. Tokyo Sect. 1A Math. **31**, 223–245 (1984)
3. Maruyama, G.: On the transition probability functions of the Markov processes. Nat. Sci. Rep. Ochanomizu Univ. **5**, 10–20 (1954)
4. McKean, H.P.: A class of Markov processes associated with non-linear parabolic equations. Proc. Natl. Acad. Sci. **56**, 1907–1911 (1966)
5. McKean, H.P.: Propagation of chaos for a class of non-linear parabolic equations. Lecture Series in Differential Equations. Catholic Univ. 41–57 (1967), Washington D.C.

6. Nagasawa, M.: Segregation of a population in an environment. J. Math. Biol. 9, 213-235 (1980)
7. Nagasawa, M.: An application of segregation model for septation of *Escherichia coli*. J. Theor. Biol. **90**, 445–455 (1981)
8. Nagasawa, M., K. Yasue: A statistical model of mesons. Publication de l'Institute recherche Mathematique Avanceé (CNRS) **33**, 1–48. Université Louis Pasteur (Strasbourg) (1982/83)
9. Nagasawa, M., H. Tanaka: A diffusion process in a singular mean-drift-field. Z. Wahrscheinlichkeitstheor. Verw. Geb. **68**, 247–269 (1985)
10. Shiga, T., H. Tanaka: Central limit theorem for a system of Markovian particles with mean field interaction. Z. Wahrscheinlichkeitstheor. Verw. Geb. **69**, 439-459 (1985)
11. Sznitman, A.S.: Non-linear reflecting diffusion processes, and propagation of chaos, and fluctuations associated. J. Funct. Anal. **56**, 311-336 (1984)
12. Tanaka, H.: Stochastic differential equations with reflecting boundary condition in convex regions. Hiroshima Math. J. 9, 163–177 (1979)
13. Tanaka, H.: Limit theorems for certain diffusion processes with interaction. In: Stochastic Analysis, pp. 469-488. Itô, K. (ed.). Tokyo: Kinokuniya Co. Ltd., Amsterdam-Oxford-New York: North-Holland Pub. Co. 1984

Received November 29, 1984

Saisho, Y. and Tanaka, H.
Osaka J. Math.
23 (1986), 725–740

STOCHASTIC DIFFERENTIAL EQUATIONS FOR MUTUALLY REFLECTING BROWNIAN BALLS

YASUMASA SAISHO AND HIROSHI TANAKA

(Received August 1, 1985)

Introduction

In this paper we construct a random motion of mutually reflecting hard balls of diameter ρ in $\boldsymbol{R}^d$ by solving certain stochastic differential equation (abbreviated: SDE) with a kind of singular drift. For simplicity we first consider the motion of mutually reflecting Brownian balls of diameter ρ. In order to construct such a motion we pose the following problem. Let W denote the space of continuous paths in $\boldsymbol{R}^d$. *Given* $w_1, \cdots, w_n \in W$ *satisfying*

(1) $$|w_i(0)-w_j(0)| \geqq \rho, \quad 1 \leqq i<j \leqq n,$$

solve the equation

(2) $$\xi_i(t) = w_i(t)+\sum_{j=1(\neq i)}^{n} \int_0^t (\xi_i(s)-\xi_j(s))\, d\phi_{ij}(s), \quad 1 \leqq i \leqq n,$$

under the following conditions (3) *and* (4).

(3) $\xi_i \in W$, $1 \leqq i \leqq n$, and $|\xi_i(t)-\xi_j(t)| \geqq \rho$, $1 \leqq i<j \leqq n$, $t \geqq 0$.

(4) ϕ_{ij}'s are continuous non-decreasing functions with $\phi_{ij}(0)=0$, $\phi_{ij}(t)=\phi_{ji}(t)$ and

$$\phi_{ij}(t) = \int_0^t \mathbf{1}_\rho(|\xi_i(s)-\xi_j(s)|)\, d\phi_{ij}(s),$$

where $\mathbf{1}_\rho(r) = 1$ if $r = \rho$, and $=0$ if $r \neq \rho$.

A pair (ξ, ϕ) of functions or simply a function ξ is called a solution of (2) provided that (2), (3) and (4) are satisfied. One of the main results in this paper is that *there exists a unique solution of* (2) for given $w_1, \cdots, w_n$. By taking $w_1, \cdots, w_n$ to be independent d-dimensional Brownian motions satisfying (1), we obtain a process $(\xi_1(t), \cdots, \xi_n(t))$. This is what we call the motion of mutually reflecting Brownian balls. $\xi_i(t)$ denotes the center of the i-th Brownian ball at time t. In analogy with Skorohod's equation for a 1-dimensional reflecting Brownian motion ([3] [5] [7]), the equation (2) may be regarded as Skorohod's equation for the motion of mutually reflecting Brownian balls.

We next consider the SDE on a probability space $(\Omega, \mathcal{F}, P)$:

$$(5) \qquad dX_i(t) = \sigma(X_i(t))\, dB_i(t) + b(X_i(t))\, dt + \sum_{j=1(\neq i)}^{n} (X_i(t) - X_j(t))\, d\phi_{ij}(t), \quad 1 \leqq i \leqq n,$$

where

$$\sigma: \boldsymbol{R}^d \to \boldsymbol{R}^d \otimes \boldsymbol{R}^d, \quad b: \boldsymbol{R}^d \to \boldsymbol{R}^d$$

are given, $X_i(0)$'s are $\mathcal{F}_0$-measurable initial values satisfying $|X_i(0) - X_j(0)| \geqq \rho$, $1 \leqq i < j \leqq n$, and $B_i(t)$, $1 \leqq i \leqq n$, are independent d-dimensional $\mathcal{F}_t$-adapted Brownian motion with $B_i(0) = 0$. Here $\{\mathcal{F}_t\}_{t \geqq 0}$ is a right continuous filtration on $(\Omega, \mathcal{F}, P)$ such that each $\mathcal{F}_t$ contains all P-negligible sets. As in (2), $X_i(t)$ and $\Phi_{ij}(t)$ should be found under the following conditions (6) and (7).

(6) $X_i(t)$'s are $\mathcal{F}_t$-adapted continuous processes with $|X_i(t) - X_j(t)| \geqq \rho$, $1 \leqq i < j \leqq n$, $t \geqq 0$.

(7) $\Phi_{ij}(t)$'s are $\mathcal{F}_t$-adapted continuous non-decreasing processes with $\Phi_{ij}(0) = 0$, $\Phi_{ij}(t) = \Phi_{ji}(t)$ and

$$\Phi_{ij}(t) = \int_0^t \mathbf{1}_\rho(|X_i(s) - X_j(s)|)\, d\Phi_{ij}(s).$$

The equation (5) may be considered as Skorohod's SDE for mutually reflecting diffusion balls with coefficients σ and b. Another main result of this paper is that *there exists a unique solution of* (5) *provided that* σ *and* b *are bounded and Lipschitz continuous.*

Our method for solving (2) and (5) is to make use of the results of [6] concerning Skorohod's equation for general domain. Skorohod's equation for a multi-dimensional domain D with reflecting boundary (a precise formulation is explained in **1**) was discussed by Tanaka [8] when D is a convex domain and then by Lions and Sznitman [4] when D is a general domain satisfying Conditions (A) and (B) (see **1**) together with the additional condition that D is *admissible*, which means roughly that D can be approximated in some sense by smooth domains. Recently, Frankowska [2] and Saisho [6] amplified Lions and Sznitman's result by removing the additional condition. Now the present discussion is based on the fact that Skorohod's equation (2) (or SDE (5)) is equivalent to Skorohod's equation (or SDE) for the domain

$$(8) \qquad D = \{(x_1, \cdots, x_n) \in \boldsymbol{R}^{nd} : |x_i - x_j| > \rho,\ 1 \leqq i < j \leqq n\}.$$

So the crucial point of our discussions is to prove that the domain D of (8) satisfies Conditions (A) and (B). In solving Skorohod's SDE for D we make use of Theorem 5.1 of Saisho [6].

In **1** we state briefly the results of Lions and Sznitman [4] and Saisho [6] concerning Skorohod's equation for a general domain. We prove that the domain D of (8) satisfies Condition (B) in **2** and (A) in **3**. We solve Skorohod's equation (2) in **4** and SDE (5) in **5**.

1. Some known results on Skorohod's equation for an N-dimensional domain with reflecting boundary

Let D be a domain in $\boldsymbol{R}^N$ and define the set $\mathcal{N}_x$ of inward normal unit vectors at $x\in\partial D$ by

$$\mathcal{N}_x = \bigcup_{r>0} \mathcal{N}_{x,r},$$
$$\mathcal{N}_{x,r} = \{\boldsymbol{n}\in\boldsymbol{R}^N\colon |\boldsymbol{n}|=1,\ B(x-r\boldsymbol{n}, r)\cap D=\phi\},$$

where $B(z, r)=\{y\in\boldsymbol{R}^N\colon |y-z|<r\}$, $z\in\boldsymbol{R}^N$. In general it can happen that $\mathcal{N}_x=\phi$. In what follows $\langle\cdot,\cdot\rangle$ denotes the usual inner product in $\boldsymbol{R}^N$. We introduce two conditions on the domain D.

Condition (*A*) (uniform exterior sphere condition). There exists a constant $r_0>0$ such that

$$\mathcal{N}_x = \mathcal{N}_{x,r_0}\neq\phi \quad \text{for any} \quad x\in\partial D.$$

Condition (*B*). There exist constants $\delta>0$ and $\beta(1\leqq\beta<\infty)$ with the following property: for any $x\in\partial D$ there exists a unit vector $\boldsymbol{l}_x$ such that

$$\langle \boldsymbol{l}_x, \boldsymbol{n}\rangle\geqq 1/\beta \quad \text{for any} \quad \boldsymbol{n}\in \bigcup_{y\in B(x,\delta)\cap\partial D} \mathcal{N}_y.$$

REMARK 1.1. For any fixed $r>0$ and a unit vector $\boldsymbol{n}$ the following two statements are equivalent.

(i) $B(x-r\boldsymbol{n}, r)\cap D=\phi$.

(ii) $\langle y-x, \boldsymbol{n}\rangle+\dfrac{1}{2r}|y-x|^2\geqq 0$ for any $y\in\bar{D}$.

REMARK 1.2. D satisfies Condition (B) if it satisfies the following condition.

Condition (*B'*). There exist $\delta>0$ and $\alpha(0\leqq\alpha<1)$ with the following property: for any $x\in\partial D$ there exists a unit vector $\boldsymbol{l}_x$ such that

$$C(y, \boldsymbol{l}_x, \alpha)\cap B(x, \delta)\subset\bar{D}, \quad \forall y\in B(x, \delta)\cap\partial D,$$

where $C(y, \boldsymbol{l}_x, \alpha)$ is the convex cone with vertex y, defined by

$$C(\boldsymbol{y}, \boldsymbol{l}_x, \alpha) = \{z \in \boldsymbol{R}^N : \langle z-\boldsymbol{y}, \boldsymbol{l}_x \rangle \geqq \alpha |z-\boldsymbol{y}|\} .$$

Denote by $W(\boldsymbol{R}^N)$ (resp. $W(\bar{D})$) the space of continuous paths in $\boldsymbol{R}^N$ (resp. $\bar{D}$). Skorohod's equation for D with reflecting boundary is written in the form

$$\xi(t) = w(t) + \int_0^t \boldsymbol{n}(s)\, d\phi(s) , \tag{1.1}$$

where $w \in W(\boldsymbol{R}^N)$ is given and satisfies $w(0) \in \bar{D}$; a solution (ξ, ϕ) of (1.1) should be found under the following conditions.

(1.2) $\xi \in W(\bar{D})$.

(1.3) ϕ is a continuous non-decreasing function such that $\phi(0) = 0$ and

$$\phi(t) = \int_0^t \mathbf{1}_{\partial D}(\xi(s))\, d\phi(s) .$$

(1.4) $\boldsymbol{n}(s) \in \mathcal{N}_{\xi(s)}$ if $\xi(s) \in \partial D$.

The following theorem was proved by Lions and Sznitman [4] under the additional condition that D is admissible. Frankowska [2] and Saisho [6] removed this additional condition. Frankowska's result is of a general type but contains what we need only in a less explicit form, so we state the theorem in the form of Saisho [6].

Theorem 1.1. *If the domain D satisfies Conditions (A) and (B), then there exists a unique solution of* (1.1) *for any given $w \in W(\boldsymbol{R}^N)$ with $w(0) \in \bar{D}$.*

Next, given

$$\sigma : \bar{D} \to \boldsymbol{R}^N \otimes \boldsymbol{R}^N, \quad b : \bar{D} \to \boldsymbol{R}^N ,$$

we consider Skorohod's SDE

$$dX(t) = \sigma(X(t))\, dB(t) + b(X(t))\, dt + \boldsymbol{n}(t)\, d\Phi(t) , \tag{1.5}$$

where the initial value $X(0) \in \bar{D}$ is assumed to be $\mathcal{F}_0$-measurable and $B(t)$ is an N-dimensional $\mathcal{F}_t$-adapted Brownian motion with $B(0)=0$. Here $\{\mathcal{F}_t\}$ is a right continuous filtration on $(\Omega, \mathcal{F}, P)$ such that $\mathcal{F}_0$ contains all P-negligible sets. A solution $(X(t), \Phi(t))$ should be found under the following conditions (1.6)–(1.8).

(1.6) $X(t)$ is a $\bar{D}$-valued $\mathcal{F}_t$-adapted continuous process.

(1.7) $\Phi(t)$ is a continuous non-decreasing process with $\Phi(0)=0$ and

$$\Phi(t) = \int_0^t \mathbf{1}_{\partial D}(X(s))\, d\Phi(s) .$$

(1.8) $\boldsymbol{n}(s)\in\mathcal{N}_{X(s)}$ if $X(s)\in\partial D$.

In addition to Conditions (A) and (B), Lions and Sznitman [4] introduced the following Condition (C) and discussed the existence and uniqueness of the solution of (1.5).

Condition (*C*). There exists a function f in $C^2(\boldsymbol{R}^N)$ which is bounded together with its first and second partial derivatives such that $\exists\gamma>0$, $\forall\boldsymbol{x}\in\partial D$, $\forall\boldsymbol{y}\in\bar{D}$, $\forall\boldsymbol{n}\in\mathcal{N}_{\boldsymbol{x}}$

$$\langle\boldsymbol{y}-\boldsymbol{x},\boldsymbol{n}\rangle+\frac{1}{\gamma}\langle\nabla f(\boldsymbol{x}),\boldsymbol{n}\rangle|\boldsymbol{y}-\boldsymbol{x}|^2\geqq 0\,.$$

The following theorem was proved by Lions and Sznitman ([4: Theorem 3.1]) under Condition (C) and the admissibility of D; however, recently Saisho ([6: Theorem 5.1]) removed these additional conditions.

Theorem 1.2 ([6]). *Let D satisfy Conditions* (*A*), (*B*) *and assume that σ and b are bounded and Lipschitz continuous. Then for any initial value $X(0)\in\bar{D}$ there exists a unique* (*strong*) *solution of* (1.5).

2. D satisfies Condition (B')

Let D be the domain in $\boldsymbol{R}^{nd}$ defined by (8). We are going to prove the following proposition.

Proposition 2.1. *The domain D satisfies Condition* (B'), *that is, there exist constants $\delta>0$ and $\alpha(0\leqq\alpha<1)$ with the following property: for any $\boldsymbol{x}\in\partial D$ there exists a unit vector $\boldsymbol{l}_{\boldsymbol{x}}$ such that*

(2.1) $$C(\boldsymbol{y},\boldsymbol{l}_{\boldsymbol{x}},\alpha)\cap B(\boldsymbol{x},\delta)\subset\bar{D}$$

holds for any $\boldsymbol{y}\in B(\boldsymbol{x},\delta)\cap\partial D$.

For a non-empty subset I of $\{1,2,\cdots,n\}$ and $\boldsymbol{x}=(x_1,\cdots,x_n)\in\bar{D}$ we set

$$\boldsymbol{x}(I)=\{x_i\colon i\in I\}\,.$$

DEFINITION 2.1. Let I, I' be non-empty subsets of $\{1,2,\cdots,n\}$.
(i) $\boldsymbol{x}(I)$ and $\boldsymbol{x}(I')$ are said to be *separated* if

(2.2) $$I\cap I'=\phi \text{ and } |x_i-x_j|\geqq 2\rho,\ \forall i\in I,\ \forall j\in I'\,.$$

(ii) $\boldsymbol{x}(I)$ is called a *cluster* if

(2.3) for any $i,j\in I$ with $i\neq j$ there exist $i_0(=i)$, $i_1,\cdots,i_{p-1},i_p(=j)$ in I such that $|x_{i_{k-1}}-x_{i_k}|<2\rho$, $1\leqq k\leqq p$.

REMARK 2.1. If $\boldsymbol{x}(I)$ is a cluster, then

$$|x_i - x_j| < 2(n-1)\rho, \quad \forall i, j \in I, \tag{2.4}$$

$$\begin{aligned} |x_i - x_I| &= \frac{1}{\#I} \Big| \sum_{j \in I} (x_i - x_j) \Big| \\ &< 2(n-1)\rho, \quad \forall i \in I, \end{aligned} \tag{2.5}$$

where $x_I = (\#I)^{-1} \sum_{j \in I} x_j$ and $\#I$ is the number of elements in I.

Let $\boldsymbol{x} = (x_1, \cdots, x_n)$. Then $\{x_1, \cdots, x_n\}$ can be represented as the sum of mutually separated clusters:

$$\{x_1, \cdots, x_n\} = \bigcup_{k=1}^{m} \boldsymbol{x}(I_k). \tag{2.6}$$

Here $1 \leqq m \leqq n$. In what follows we assume that $\boldsymbol{x} \in \partial D$ and keep it fixed, so $I_1, \cdots, I_m$ appearing in (2.6) are also fixed. Let $c > 1$ be a constant which will be determined later and set

$$\begin{cases} u_i = x_{I_k} + c(x_i - x_{I_k}), & i \in I_k,\ 1 \leqq k \leqq m, \\ \boldsymbol{u} = (u_1, \cdots, u_n) \in \boldsymbol{R}^{nd}. \end{cases}$$

Also we set for $\boldsymbol{y} = (y_1, \cdots, y_n) \in \boldsymbol{R}^{nd}$

$$\begin{cases} v_i = y_{I_k} + c(y_i - y_{I_k}), & i \in I_k,\ 1 \leqq k \leqq m, \\ \boldsymbol{v} = (v_1, \cdots, v_n) \in \boldsymbol{R}^{nd}. \end{cases}$$

We are going to prove Proposition 2.1 with

$$\begin{cases} \boldsymbol{l}_x = \dfrac{\boldsymbol{u} - \boldsymbol{x}}{|\boldsymbol{u} - \boldsymbol{x}|}, \\ \delta = \dfrac{\varepsilon}{4(c-1)+2}, \\ \alpha = \cos\theta, \end{cases} \tag{2.7}$$

where c, ε and θ are constants determined by

$$\begin{cases} (c-1)(n-1) = 1/8, \\ \varepsilon = (c-1)\rho/4, \end{cases} \tag{2.8}$$

$$\sin\frac{\theta}{2} = \frac{\varepsilon}{8(c-1)\sqrt{n}\,(n-1)\rho}, \quad 0 < \theta \leqq \pi/2. \tag{2.9}$$

From now on let c, ε, δ be constants determined by (2.7) and (2.8). Let $\boldsymbol{y} \in B(\boldsymbol{x}, \delta) \cap \partial D$ and set

$$\boldsymbol{z} = \boldsymbol{y} + (\boldsymbol{u} - \boldsymbol{x}).$$

Lemma 2.1. $|\boldsymbol{u}-\boldsymbol{x}|=|\boldsymbol{z}-\boldsymbol{y}|<2(c-1)\sqrt{n}(n-1)\rho$.

Proof. Since $i\in I_k$ for some k, we have

$$|u_i-x_i| = (c-1)|x_i-x_{I_k}|<2(c-1)(n-1)\rho \tag{2.10}$$

by (2.5), and hence we obtain the lemma.

Lemma 2.2. (i) *For any* $\boldsymbol{u}'\in B(\boldsymbol{u},\varepsilon)$, $\boldsymbol{v}'\in B(\boldsymbol{v},\varepsilon)$ *and* $0<t\leqq 1$,

$$\boldsymbol{u}'' = (1-t)\boldsymbol{x}+t\boldsymbol{u}'\in D, \quad \boldsymbol{v}'' = (1-t)\boldsymbol{y}+t\boldsymbol{v}'\in D.$$

(ii) $B(\boldsymbol{u},\varepsilon)\subset D$, $B(\boldsymbol{v},\varepsilon)\subset D$.

Proof. Since (ii) can be proved by setting $t=1$ in (i), it is enough to prove (i), that is,

$$|u_i''-u_j''|>\rho, \quad |v_i''-v_j''|>\rho, \quad 1\leqq i<j\leqq n.$$

Case (I): $i,j\in I_k$ *for some* k. Since

$$\begin{aligned} u_i''-u_j'' &= (1-t)(x_i-x_j)+t\{(u_i-u_j)+(u_i'-u_i)-(u_j'-u_j)\} \\ &= \{1+(c-1)t\}(x_i-x_j)+t\{(u_i'-u_i)-(u_j'-u_j)\}, \end{aligned}$$

we have

$$|u_i''-u_j''|\geqq\{1+(c-1)t\}\rho-2\varepsilon t>\rho,$$

and a similar inequality for $|v_i''-v_j''|$.
Case (II): $i\in I_k$, $j\in I_l$ $(k\neq l)$. Since

$$u_i''-x_i = t(u_i'-x_i) = t(u_i-x_i)+t(u_i'-u_i),$$

we have

$$\begin{aligned} |u_i''-x_i| &\leqq |u_i-x_i|+|u_i'-u_i| \quad (\text{setting } t=1) \\ &<2(c-1)(n-1)\rho+\varepsilon \quad (\text{by } (2.10)), \end{aligned}$$

and hence, making use of the inequality $|x_i-x_j|\geqq 2\rho$ which is a consequence of the assumption that $\boldsymbol{x}(I_k)$ and $\boldsymbol{x}(I_l)$ are separated, we have

$$\begin{aligned} |u_i''-u_j''| &\geqq |x_i-x_j|-|u_i''-x_i|-|u_j''-x_j| \\ &>2\rho-4(c-1)(n-1)\rho-2\varepsilon>\rho. \end{aligned}$$

Next, to prove the inequality for $|v_i''-v_j''|$, we notice that

$$\begin{aligned} |(y_i-y_{I_k})-(x_i-x_{I_k})| &= \left|(y_i-x_i)-\frac{1}{\# I_k}\sum_{p\in I_k}(y_p-x_p)\right| \\ &\leqq 2\max_{1\leqq p\leqq n}|y_p-x_p|\leqq 2|\boldsymbol{y}-\boldsymbol{x}|<2\delta, \end{aligned}$$

from which it follows that

$$|y_i-y_{I_k}|<|x_i-x_{I_k}|+2\delta<2(n-1)\rho+2\delta \quad (\text{use } (2.5)).$$

Therefore

$$|v_i-y_i| = (c-1)|y_i-y_{I_k}|<2(c-1)(n-1)\rho+2(c-1)\delta,$$

and hence

$$|v_i''-y_i|\leqq t|v_i-y_i|+t|v_i'-v_i|<2(c-1)(n-1)\rho+2(c-1)\delta+\varepsilon.$$

Thus we have

$$\begin{aligned}|v_i''-v_j''| &\geqq |y_i-y_j|-|v_i''-y_i|-|v_j''-y_j|\\ &\geqq |x_i-x_j|-|y_i-x_i|-|y_j-x_j|-|v_i''-y_i|-|v_j''-y_j|\\ &>2\rho-2\delta-4(c-1)(n-1)\rho-4(c-1)\delta-2\varepsilon\\ &>\rho.\end{aligned}$$

The proof of the lemma is finished.

Lemma 2.3. $|z-v|<\varepsilon/2$.

Proof. Taking I_k such that $i\in I_k$, we have

$$\begin{aligned}|z_i-v_i| &= |(c-1)(x_i-y_i)-(c-1)(x_{I_k}-y_{I_k})|\\ &\leqq (c-1)|x_i-y_i|+(c-1)|x_{I_k}-y_{I_k}|,\end{aligned}$$

and hence

$$(2.11)\qquad \begin{aligned}|\boldsymbol{z}-\boldsymbol{v}|^2 &= \sum_{i=1}^{n}|z_i-v_i|^2\\ &\leqq 2(c-1)^2|\boldsymbol{x}-\boldsymbol{y}|^2+2(c-1)^2\sum_{k=1}^{m}\#I_k|x_{I_k}-y_{I_k}|^2.\end{aligned}$$

Because $|x_{I_k}-y_{I_k}|^2\leqq(\#I_k)^{-1}\sum_{j\in I_k}|x_j-y_j|^2$, we have

$$\sum_{k=1}^{m}\#I_k|x_{I_k}-y_{I_k}|^2\leqq|\boldsymbol{x}-\boldsymbol{y}|^2,$$

and (2.11) yields

$$|\boldsymbol{z}-\boldsymbol{v}|\leqq 2(c-1)|\boldsymbol{x}-\boldsymbol{y}|<2(c-1)\delta<\varepsilon/2.$$

The proof of the lemma is finished.

Proof of Proposition 2.1. By (ii) of Lemma 2.2 and Lemma 2.3 we have

$$B(\boldsymbol{z},\varepsilon/2)\subset B(\boldsymbol{v},\varepsilon)\subset D,$$

and hence by (i) of Lemma 2.2

(2.12) $\bar{D}\supset$the convex hull Γ of the set $B(z, \varepsilon/2)\cup\{y\}$.

We have also

(2.13) $$|z-y| = |u-x|\geqq\varepsilon>2\delta,$$

because $x\notin B(u, \varepsilon)$ by (ii) of Lemma 2.2. Therefore, if θ is defined by (2.9), then (2.12) combined with Lemma 2.1 and (2.13) implies that

(2.14) $$C(y, l_x, \cos\theta)\cap B(y, 2\delta)\subset\Gamma\subset\bar{D}.$$

Since $B(x, \delta)\subset B(y, 2\delta)$, (2.14) implies (2.1) with $\alpha=\cos\theta$. The proof of Proposition 2.1 is finished.

3. *D* satisfies Condition (A)

For $1\leqq i<j\leqq n$ let D_{ij} be the domain defined by

$$D_{ij} = \{x = (x_1, \cdots, x_n)\in \boldsymbol{R}^{nd}: |x_i-x_j|>\rho\}.$$

The domain D of (8) is expressed as

$$D = \bigcap_{1\leqq i<j\leqq n} D_{ij}$$

and $x\in\partial D$ implies $x\in\bigcap_{(i,j)\in L}\partial D_{ij}$, where

$$L = L_x = \{(i,j): 1\leqq i<j\leqq n,\ |x_i-x_j| = \rho\}.$$

Obviously, if $x\in\bar{D}$ then $x\in\partial D$ is equivalent to $L_x\neq\phi$. For $(i,j)\in L_x$ we define a unit vector n_{ij} in $\boldsymbol{R}^{nd}$ by

$$n_{ij} = \{0, \cdots, 0, \underset{(i\text{-th})}{\frac{x_i-x_j}{\sqrt{2}\rho}}, 0, \cdots, 0, \underset{(j\text{-th})}{\frac{x_j-x_i}{\sqrt{2}\rho}}, 0, \cdots, 0\},$$

let $\alpha(0\leqq\alpha<1)$ be the constant appearing in Proposition 2.1 and set

$$r_0 = \rho\sqrt{\frac{1-\alpha^2}{2}}.$$

Proposition 3.1. *The domain D satisfies Condition (A) with* $r_0=\rho\sqrt{\frac{1-\alpha^2}{2}}$, *and for any* $x\in\partial D$

(3.1) $$\mathcal{N}_x = \{n = \sum_{(i,j)\in L} c_{ij}n_{ij}: c_{ij}\geqq 0, |n| = 1\}.$$

Lemma 3.1. (i) $B(x-2^{-1/2}\rho n_{ij}, 2^{-1/2}\rho)\cap D_{ij}=\phi$, $x\in\partial D_{ij}$.
(ii) $B(x-2^{-1/2}\rho n_{ij}, 2^{-1/2}\rho)\cap D=\phi$ *for any* $(i,j)\in L_x$, $x\in\partial D$, *that is,* $n_{ij}\in$

$\mathcal{N}_{x,2^{-1/2}\rho}$ *for any* $(i,j)\in L_x$.

Proof. Since (ii) follows immediately from (i), it is enough to prove (i). We set

$$y = x - 2^{-1/2}\rho n_{ij},$$

that is,

$$y_k = \begin{cases} (x_i+x_j)/2 & \text{for} \quad k=i,j, \\ x_k & \text{for} \quad k \neq i,j. \end{cases}$$

Then for any $z\in B(y, 2^{-1/2}\rho)$ we have

$$\begin{aligned} |z_i-z_j| &\leqq |z_i-y_i|+|y_i-y_j|+|y_j-z_j| \\ &= |z_i-y_i|+|z_j-y_j| \leqq \sqrt{2}\,|z-y| < \rho, \end{aligned}$$

and hence $z\notin D_{ij}$, completing the proof.

Lemma 3.2. *Let* $\mathcal{N}'_x$ *be the right-hand side of* (3.1). *Then*

$$\mathcal{N}'_x \subset \mathcal{N}_{x,r_0}, \quad x\in\partial D.$$

Proof. By Remark 1.1 it is enough to prove that for any $n\in\mathcal{N}'_x$

$$\langle y-x, n\rangle + \frac{1}{2r_0}|x-y|^2 \geqq 0, \quad \forall y\in\bar{D}.$$

Let $l=l_x$ be the unit vector appearing in Proposition 2.1. Then by Proposition 2.1

$$C(x, l, \alpha)\cap B(x, \delta)\subset\bar{D}$$

from which it follows that

$$x-\mathcal{N}_x\subset C(x, l, \alpha)^*, \tag{3.2}$$

where $C(x, l, \alpha)^*$ is the *dual cone* of $C(x, l, \alpha)$, that is,

$$\begin{aligned} C(x, l, \alpha)^* &= \{z\in R^{nd}: \langle z-x, y-x\rangle\leqq 0, \ \forall y\in C(x, l, \alpha)\} \\ &= \{z\in R^{nd}: \langle z-x, -l\rangle \geqq \sqrt{1-\alpha^2}\,|z-x|\}. \end{aligned}$$

From (ii) of Lemma 3.1 and (3.2)

$$x-n_{ij}\in C(x, l, \alpha)^*, \quad \forall(i,j)\in L_x,$$

and hence

$$\langle n_{ij}, l\rangle \geqq \sqrt{1-\alpha^2}, \quad \forall(i,j)\in L_x. \tag{3.3}$$

Now let $\boldsymbol{n}\in\mathcal{N}'_x$ be expressed as

$$\boldsymbol{n}=\sum_{(i,j)\in L} c_{ij}\boldsymbol{n}_{ij},\quad c_{ij}\geqq 0\,.$$

Then by (3.3)

$$1\geqq\langle\boldsymbol{n},\boldsymbol{l}\rangle=\sum_{(i,j)\in L} c_{ij}\langle\boldsymbol{n}_{ij},\boldsymbol{l}\rangle\geqq\sqrt{1-\alpha^2}\sum_{(i,j)\in L} c_{ij}$$

and hence

$$\frac{1}{2r_0}=\frac{1}{\sqrt{2}\,\rho\sqrt{1-\alpha^2}}\geqq\frac{1}{\sqrt{2}\,\rho}\sum_{(i,j)\in L} c_{ij}\,.$$

Therefore, for any $\boldsymbol{y}\in\bar{D}$

$$\begin{aligned}&\langle\boldsymbol{y}-\boldsymbol{x},\boldsymbol{n}\rangle+\frac{1}{2r_0}|\boldsymbol{x}-\boldsymbol{y}|^2\\ &\geqq\sum_{(i,j)\in L} c_{ij}\langle\boldsymbol{y}-\boldsymbol{x},\boldsymbol{n}_{ij}\rangle+\frac{1}{\sqrt{2}\,\rho}\sum_{(i,j)\in L} c_{ij}|\boldsymbol{x}-\boldsymbol{y}|^2\\ &=\sum_{(i,j)\in L} c_{ij}\{\langle\boldsymbol{y}-\boldsymbol{x},\boldsymbol{n}_{ij}\rangle+\frac{1}{\sqrt{2}\,\rho}|\boldsymbol{x}-\boldsymbol{y}|^2\}\geqq 0\end{aligned}$$

by (ii) of Lemma 3.1 and Remark 1.1. The proof of Lemma 3.2 is finished.

Lemma 3.3. *For any $\varepsilon(0<\varepsilon<1)$ and $\boldsymbol{x}\in\partial D$ there exists $\delta'>0$ such that*

$$\{\bigcap_{(i,j)\in L} C_{ij}(\boldsymbol{x},\varepsilon)\}\cap B(\boldsymbol{x},\delta')\subset D\cup\{\boldsymbol{x}\}\,,$$

where

$$C_{ij}(\boldsymbol{x},\varepsilon)=\{\boldsymbol{y}\in\boldsymbol{R}^{nd}\colon\langle\boldsymbol{y}-\boldsymbol{x},\boldsymbol{n}_{ij}\rangle\geqq\varepsilon|\boldsymbol{y}-\boldsymbol{x}|\},\ (i,j)\in L\,.$$

Proof. Let $\boldsymbol{y}\in C_{ij}(\boldsymbol{y},\varepsilon)$, $\boldsymbol{y}\neq\boldsymbol{x}$. Then $\boldsymbol{y}$ can be expressed as

$$\boldsymbol{y}=\boldsymbol{x}+\boldsymbol{z},\quad\langle\boldsymbol{z},\boldsymbol{n}_{ij}\rangle\geqq\varepsilon|\boldsymbol{z}|>0\,,$$

and hence

$$\begin{aligned}|y_i-y_j|^2&=|x_i-x_j|^2+2\langle x_i-x_j,z_i-z_j\rangle+|z_i-z_j|^2\\ &\geqq\rho^2+2\langle z_i-z_j,x_i-x_j\rangle>\rho^2\,,\end{aligned}$$

because

$$\begin{aligned}0<\langle\boldsymbol{z},\boldsymbol{n}_{ij}\rangle&=\frac{1}{\sqrt{2}\,\rho}\langle z_i,x_i-x_j\rangle+\frac{1}{\sqrt{2}\,\rho}\langle z_j,x_j-x_i\rangle\\ &=\frac{1}{\sqrt{2}\,\rho}\langle z_i-z_j,x_i-x_j\rangle\,.\end{aligned}$$

Therefore $C_{ij}(x, \varepsilon) \subset D_{ij} \cup \{x\}$, and hence

$$\bigcap_{(i,j)\in L} C_{ij}(x, \varepsilon) \subset \{\bigcap_{(i,j)\in L} D_{ij}\} \cup \{x\} . \tag{3.4}$$

Since $|x_i - x_j| > \rho$ for any $(i,j) \notin L_x$, there exists $\delta' > 0$ such that for any $y \in B(x, \delta')$

$$|y_i - y_j| > \rho \quad \text{for any} \quad (i,j) \notin L_x .$$

This combined with (3.4) implies

$$\{\bigcap_{(i,j)\in L} C_{ij}(x, \varepsilon)\} \cap B(x, \delta') \subset D \cup \{x\} .$$

The proof of the lemma is finished.

Proof of Proposition 3.1. We make use of Lemma 3.3 and then Corollary in [1: p. 11] to obtain

$$x - \mathcal{N}_x \subset \{\bigcap_{(i,j)\in L} C_{ij}(x, \varepsilon)\}^* = \overline{\sum_{(i,j)\in L} C_{ij}(x, \varepsilon)^*} ,$$

where Σ means the vector sum. Thus

$$x - \mathcal{N}_x \subset \bigcap_{0<\varepsilon<1} \{\overline{\sum_{(i,j)\in L} C_{ij}(x, \varepsilon)^*}\} .$$

But the right-hand side of the above is the convex cone with vertex x spanned by $\{x - n_{ij}, (i,j) \in L_x\}$, so we have

$$\mathcal{N}_x \subset \mathcal{N}'_x ,$$

which combined with Lemma 3.2 implies

$$\mathcal{N}_x = \mathcal{N}_{x,r_0} = \mathcal{N}'_x .$$

This means that Proposition 3.1 holds.

4. Mutually reflecting Brownian balls

Since the domain D satisfies Conditions (A) and (B′) (and hence (B) by Remark 1.2), Theorem 1.1 guarantees the existence and uniqueness of the solution of Skorohod's equation for D:

$$\xi(t) = w(t) + \int_0^t n(s)\, d\phi(s) , \tag{4.1}$$

where $w = (w_1, \cdots, w_n)$, $w_k \in W$, $1 \leqq k \leqq n$, and $|w_i(0) - w_j(0)| \geqq \rho$, $1 \leqq i < j \leqq n$. A solution of (4.1) is a pair (ξ, ϕ) of functions satisfying (4.1), (1.2), (1.3) and (1.4). The component-wise expression of (4.1) is

(4.2) $$\xi_k(t)=w_k(t)+\int_0^t n_k(s)\,d\phi(s),\quad 1\leqq k\leqq n\,.$$

In this section we prove the following theorem by showing that (4.2) is equivalent to the equation

(2) $$\xi_k(t)=w_k(t)+\sum_{j=1(\neq k)}^{n}\int_0^t(\xi_k(s)-\xi_j(s))\,d\phi_{kj}(s),\quad 1\leqq k\leqq n\,.$$

Theorem 4.1. *There exists a unique solution of Skorohod's equation* (2) *for the motion of mutually reflecting Brownian balls.*

Proof. By Proposition 3.1 we have

(4.3a) $$\boldsymbol{n}(s)=\sum_{1\leqq i<j\leqq n}c_{ij}(s)\,\boldsymbol{n}_{ij}(\xi(s))\,,$$

(4.3b) $$\boldsymbol{n}_{ij}(\xi(s))=(0,\cdots,0,\underset{(i\text{—th})}{\frac{\xi_i(s)-\xi_j(s)}{\sqrt{2}\rho}},0,\cdots,0,\underset{(j\text{—th})}{\frac{\xi_j(s)-\xi_i(s)}{\sqrt{2}\rho}},0,\cdots,0)\,,$$

(4.3c) $$c_{ij}(s)\geqq 0,\quad 1\leqq i<j\leqq n\,,$$

(4.3d) $$c_{ij}(s)=0\quad\text{for}\quad (i,j)\notin L_{\xi(s)}\,.$$

The component-wise expression of $\boldsymbol{n}(s)$ is

$$n_k(s)=\sum_{j:j>k}c_{kj}(s)\,\frac{\xi_k(s)-\xi_j(s)}{\sqrt{2}\rho}+\sum_{i:i<k}c_{ik}(s)\,\frac{\xi_k(s)-\xi_i(s)}{\sqrt{2}\rho}\,.$$

Therefore if we define $c_{ij}(s)$ for $i>j$ by $c_{ij}(s)=c_{ji}(s)$, then we have

$$n_k(s)=\sum_{j:j\neq k}c_{kj}(s)\,\frac{\xi_k(s)-\xi_j(s)}{\sqrt{2}\rho}\,,$$

and hence

$$\xi_k(t)=w_k(t)+\sum_{j:j\neq k}\int_0^t c_{kj}(s)\,\frac{\xi_k(s)-\xi_j(s)}{\sqrt{2}\rho}\,d\phi(s)\,.$$

So if we set

$$\phi_{kj}(t)=\frac{1}{\sqrt{2}\rho}\int_0^t c_{kj}(s)\,d\phi(s)\,,$$

we have

(4.4) $$\phi_{kj}(t)=\phi_{jk}(t)\,,$$

$$\phi_{kj}(t) = \int_0^t \mathbf{1}_\rho(|\xi_k(s)-\xi_j(s)|)\, d\phi_{kj}(s)\,, \tag{4.5}$$

because $\phi_{kj}(t)$ can increase only when $|\xi_k(t)-\xi_j(t)|=\rho$ by (4.3d). Thus (4.1) finally yields (2). Recalling (3.1), we see easily that (2) also implies (4.1) with

$$n(t) = \sum_{1\leq i<j\leq n} a_{ij}(t)\, n_{ij}(\xi(t))/|\sum_{1\leq i<j\leq n} a_{ij}(t)\, n_{ij}(\xi(t))|\,,$$
$$d\phi(t) = \sqrt{2}\,\rho|\sum_{1\leq i<j\leq n} a_{ij}(t)\, n_{ij}(\xi(t))|\, d\bar{\phi}(t)\,,$$

where $\bar{\phi}(t)=\sum_{1\leq i<j\leq n}\phi_{ij}(t)$ and $a_{ij}(t)$ is the Radon-Nikodym derivative of $d\phi_{ij}(t)$ with respect to $d\bar{\phi}(t)$. The proof of Theorem 4.1 is finished.

5. Skorohod's SDE for mutually reflecting diffusion balls

Let $B_i(t)$, $1\leqq i\leqq n$, be independent d-dimensional $\mathcal{F}_t$-Brownian motions with $B_i(0)=0$ defined on a probability space $(\Omega, \mathcal{F}, P)$ with a filtration $\{\mathcal{F}_t\}_{t\geqq 0}$. We assume that each $\mathcal{F}_t$ contains all P-null sets and $\mathcal{F}_t=\bigcap_{\varepsilon>0}\mathcal{F}_{t+\varepsilon}$. Given the coefficients

$$\sigma\colon \mathbf{R}^d \to \mathbf{R}^d\otimes\mathbf{R}^d, \quad b\colon \mathbf{R}^d \to \mathbf{R}^d\,,$$

we consider the following Skorohod's SDE for mutually reflecting diffusion balls:

$$\begin{aligned} dX_i(t) &= \sigma(X_i(t))\, dB_i(t)+b(X_i(t))\, dt \\ &\quad+\sum_{j=1(\neq i)}^{n}(X_i(t)-X_j(t))\, d\Phi_{ij}(t), \quad 1\leqq i\leqq n\,, \end{aligned} \tag{5.1}$$

where the initial values are assumed to be $\mathcal{F}_0$-measurable random variables satisfying $|X_i(0)-X_j(0)|>\rho$, $1\leqq i<j\leqq n$. The solution $X_i(t)$, $1\leqq i\leqq n$, should be found under the following conditions.

(5.2) $X_i(t)$'s are $\mathcal{F}_t$-adapted continuous processes with $|X_i(t)-X_j(t)|\geqq\rho$, $1\leqq i<j\leqq n$, $t\geqq 0$.

(5.3) $\Phi_{ij}(t)$'s are $\mathcal{F}_t$-adapted continuous non-decreasing processes with $\Phi_{ij}(0)=0$, $\Phi_{ij}(t)=\Phi_{ji}(t)$ and

$$\Phi_{ij}(t) = \int_0^t \mathbf{1}_\rho(|X_i(s)-X_j(s)|)\, d\Phi_{ij}(s)\,.$$

In this section we prove the following theorem.

Theorem 5.1. *Suppose σ and b are bounded and Lipschitz continuous. Then there exists a unique strong solution of* (5.1).

Proof. Let

$$D = \{x = (x_1, \cdots, x_n) \in \boldsymbol{R}^{nd} \colon |x_i - x_j| > \rho,\ 1 \leqq i < j \leqq n\},$$

and for $x = (x_1, \cdots, x_n) \in \boldsymbol{R}^{nd}$ set

$$\boldsymbol{\sigma}(x) = \begin{bmatrix} \sigma(x_1) & & & 0 \\ & \sigma(x_2) & & \\ & & \ddots & \\ 0 & & & \sigma(x_n) \end{bmatrix}, \quad \boldsymbol{b}(x) = \begin{bmatrix} b(x_1) \\ \vdots \\ b(x_n) \end{bmatrix}.$$

Then as in 4 the SDE (5.1) can be regarded as Skorohod's SDE for D

$$(5.4) \qquad dX(t) = \boldsymbol{\sigma}(X(t))\, d\boldsymbol{B}(t) + \boldsymbol{b}(X(t))\, dt + \boldsymbol{n}(t)\, d\Phi(t),$$

where $\boldsymbol{B}(t)$ is an nd-dimensional $\mathcal{F}_t$-Brownian motion. The solution $X(t)$ is to be found under the following conditions.

(5.5) $X(t)$ is a $\bar{D}$-valued continuous process.

(5.6) $\boldsymbol{n}(t) \in \mathcal{N}_{X(t)}$ if $X(t) \in \partial D$.

(5.7) $\Phi(t)$ is a continuous non-decreasing process and

$$\Phi(t) = \int_0^t \mathbf{1}_{\partial D}(X(s))\, d\Phi(s).$$

But since D satisfies Conditions (A) and (B), it follows immediately from Theorem 5.1 in [6] that the Skorohod's SDE (5.4) has a unique strong solution. The proof is finished.

References

[1] W. Fenchel: Convex cones, sets, and functions, Lecture notes, Princeton Univ., 1953.

[2] H. Frankowska: *A viability approach to the Skorohod problem,* Stochastics **14** (1985), 227–244.

[3] N. Ikeda and S. Watanabe: Stochastic differential equations and diffusion processes, North Holland-Kodansha, Amsterdam-Tokyo, 1981.

[4] P.L. Lions and A.S. Sznitman: *Stochastic differential equations with reflecting boundary conditions,* Comm. Pure Appl. Math. **37** (1984), 511–537.

[5] H.P. McKean: Stochastic integrals, Academic Press, New York, 1969.

[6] Y. Saisho: *Stochastic differential equations for multi-dimensional domain with reflecting boundary,* to appear in Probab. Th. Rel. Fields.

[7] A.V. Skorohod: *Stochastic equations for diffusion processes in a bounded region* 1, 2, Theor. Veroyatnost. i Primenen. **6** (1961), 264–274; **7** (1962), 3–23.

[8] H. Tanaka: *Stochastic differential equations with reflecting boundary condition in*

convex regions, Hiroshima Math. J. **9** (1979), 163–177.

Department of Mathematics
Faculty of Science and Technology
Keio University
Kohokuku, Yokohama 223
Japan

LIMIT DISTRIBUTION FOR 1-DIMENSIONAL DIFFUSION IN A REFLECTED BROWNIAN MEDIUM

By H. Tanaka

Introduction

In analogy with Sinai's problem [8] on a random walk in a random medium, Brox [1] considered the diffusion process $X(t)$ described by the stochastic differential equation

$$(1) \qquad dX(t) = dB(t) - \frac{1}{2} W'(X(t))dt , \quad X(0) = 0 ,$$

where $\{W(x), x \in \mathbb{R}\}$ is a Brownian medium independent of another Brownian motion $B(t)$, and proved that $(\log t)^{-2}X(t)$ converges in distribution as $t \to \infty$. Similar results in the case of a considerably wider class of self-similar random media were obtained by Schumacher [7]. Recently Kesten [5] obtained the exact form of the limit distribution for Sinai's random walk as well as for a diffusion in a Brownian medium. See also [2] for a related problem.

In this paper we substitute $W(x)$ in (1) by a nonnegative reflected Brownian medium and find the corresponding limit distribution. The result was already anounced in [9] without proof but the Laplace transform of the limit distribution given in [9: §3] is not correct. We give here a full proof to the whole result of [9: §3] with a correction (see Theorem 1 and 2 below). Our method is similar to that of [1].

<u>Theorem 1</u>. Let $X(t)$ be a solution of (1) where $W_+ = \{W(x), x \geq 0\}$ and $W_- = \{W(-x), x \geq 0\}$ are independent reflected Brownian motions on the half line $[0, \infty)$ starting from 0 which are also independent of the Brownian motion $B(t)$. Then the distribution of $(\log t)^{-2}X(t)$ converges as $t \to \infty$ to the distribution μ defined by

$$(2) \qquad \mu = \int m_W Q(dW)$$

where m_W is the probability measure on $\mathbb{R}$ defined by (3.1) and Q is the probability measure on the space of media $\mathbb{W} = C(\mathbb{R} \to [0,\infty)) \cap \{W: W(0)=0\}$ such that $W_{\pm}$ are independent reflected Brownian motions on $[0, \infty)$.

<u>Theorem 2</u>. μ has a symmetric density and for $\lambda > 0$

$$(3) \qquad \int_0^\infty e^{-\lambda x} \mu(dx) = \int_0^\infty \frac{\sinh\sqrt{2\lambda}}{\sqrt{2\lambda}\cosh\sqrt{2\lambda} + t\sinh\sqrt{2\lambda}} \cdot \frac{dt}{(1+t)^2} .$$

The present case is not contained in the framework of [7] since the nonnegative reflected medium $W(x)$ has (uncountably) many points giving its minimum. The case of a nonpositive reflected Brownian medium was discussed in [9] . Some generalizations will be discussed in [5] .

Acknowledgment. I wish to thank Professor H.Kesten for pointing out mistakes of the first version of this paper.

§1. Preliminaries and exit times from valleys

Let $\mathbb{W}$ and Q be defined as in Theorem 1. For each $W \in \mathbb{W}$ solutions of the stochastic differential equation (1) define a diffusion process in $\mathbb{R}$ with generator

$$(1.1) \qquad \frac{1}{2} e^{W(x)} \frac{d}{dx}\left(e^{-W(x)} \frac{d}{dx}\right) .$$

Such a diffusion can be constructed from a Brownian motion $B(t)$[1)] as follows ([4]). Let Ω be the space of continuous functions $\omega : [0, \infty) \to \mathbb{R}$ with $\omega(0) = 0$, and denote by P the Wiener measure on Ω. Denote the value of ω at time t by $\omega(t)$ or by $B(t)$ and put

$$L(t,x) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int_0^t \mathbb{1}_{[x,x+\varepsilon)}(B(s))ds \qquad \text{(local time)},$$

$$S(x) = \int_0^x e^{W(y)} dy ,$$

$$A(t) = \int_0^t e^{-2W(S^{-1}(B(s)))} ds = \int_{\mathbb{R}} e^{-2W(S^{-1}(x))} L(t,x)dx , \quad t \geq 0 ,$$

S^{-1}, A^{-1} = the inverse functions .

Then the process $X(t, W) = S^{-1}(B(A^{-1}(t)))$ defined on the probability space (Ω, P) is a diffusion process with generator (1.1) starting at 0. If we set $(W^x)(\cdot) = W(\cdot + x)$, then $X^x(t, W) = x + X(t, W^x)$ is a diffusion process with generator (1.1) starting at x . Let

$$T(x_1, x_2) = \inf \{t \geq 0 : B(t) \notin (x_1, x_2)\} ,$$

1) The Brownian motion here is not the same as the one in (1) but we use the same notation $B(t)$.

$$L(x_1, x_2, x) = L(T(x_1, x_2), x) , \quad x \in \mathbb{R} ,$$

$$S_\lambda(x) = \int_0^x e^{\lambda W(y)} dy ,$$

$$X_\lambda(t) = X(t, \lambda W) , \quad X_\lambda^x(t) = x + X(t, \lambda W^x) .$$

Next we define a valley. Given $W \in \mathbb{W}$, a quartet $V = (a, b_1, b_2, c)$ is called a <u>valley</u> of W if

(i) $a < b_1 < 0 < b_2 < c$,

(ii) $W(b_1) = W(b_2) = 0, \quad W(a) = W(c) = D$,

(iii) $0 < W(x) < W(a)$ for $a < x < b_1$,

$0 < W(x) < W(c)$ for $b_2 < x < c$,

(iv) $A_- = \sup \{W(y) - W(x) : a < x < y < b_2\} < D$,

$A_+ = \sup \{W(x) - W(y) : b_1 < x < y < c\} < D$.

We call D (resp. $A = A_- \vee A_+$ [2)]) the <u>depth</u> (resp. the <u>inner directed ascent</u>) of V. It is clear that there exist valleys of W with $A < 1 < D$ for almost all reflected Brownian media W .

In what follows let $W \in \mathbb{W}$ be given and $V = (a, b_1, b_2, c)$ be a valley of W with the depth D and the inner directed ascent A . We put

$$T_\lambda^x = T_\lambda^x(a, c) = \inf \{t \geq 0 : X_\lambda^x(t) \notin (a, c)\} .$$

The following three lemmas were proved in [1] .

<u>Lemma 1</u>. For $a < x < c$

$$T_\lambda^x(a, c) \overset{d}{=} \int_a^c L(\hat{S}_\lambda(a), \hat{S}_\lambda(c), \hat{S}_\lambda(y)) e^{-\lambda W(y)} dy ,$$

where

$$\hat{S}_\lambda(y) = \int_x^y e^{\lambda W(z)} dz$$

and $\overset{d}{=}$ means the equality in distribution.

<u>Lemma 2</u>. For each $\lambda > 0$

$$\{L(\lambda x_1, \lambda x_2, \lambda x), x \in \mathbb{R}\} \overset{d}{=} (\lambda L(x_1, x_2, x), x \in \mathbb{R}\}.$$

---

2) $a \vee b = \max \{a, b\}$, $a \wedge b = \min \{a, b\}$.

Lemma 3. For $\lambda > 0$ and $W \in \mathbb{W}$

(1.2) $$\{X(t, \lambda W_\lambda), t \geq 0\} \overset{d}{=} \{\lambda^{-2}X(\lambda^4 t, W), t \geq 0\},$$

where W_λ ($\in \mathbb{W}$) is defined by

$$W_\lambda(x) = \lambda^{-1}W(\lambda^2 x), \quad x \in \mathbb{R}.$$

The following lemma plays an essential role in our discussions.

Lemma 4. For any $\lambda > 0$ and $[u, v] \subset (a, c)$

$$\inf_{u \leq x \leq v} P\left\{e^{\lambda(D-\delta)} < T_\lambda^x < e^{\lambda(D+\delta)}\right\} \to 1, \quad \lambda \to \infty.$$

Proof. The proof is similar to that of the corresponding lemma of [1] but even much simpler. Let $x \in [u, v]$ be fixed. Setting

$$s_\lambda(y) = \widehat{S}_\lambda(y)/\widehat{S}_\lambda(c) = \int_x^y e^{\lambda W(z)}dz \Big/ \int_x^c e^{\lambda W(z)}dz$$

and applying Lemma 1 and 2, we have

$$T_\lambda^x \overset{d}{=} \widehat{S}_\lambda(c)\int_a^c L(s_\lambda(a), 1, s_\lambda(y))e^{-\lambda W(y)}dy.$$

Since

$$\widehat{S}_\lambda(c) \leq (c - x)\exp\left\{\lambda \max_{[x,c]} W\right\} \quad ^{3)}$$

$$T_\lambda^x \overset{d}{\leq} (c - x)(c - a)\exp\left\{\lambda \max_{[x,c]} W - \lambda \min_{[a,c]} W\right\}L' \leq (c - a)^2 L' e^{\lambda D},$$

$$L' = \max_{y \leq 1} L(-\infty, 1, y),$$

we have

$$P\left\{T_\lambda^x > e^{\lambda(D+\delta)}\right\}$$
$$\leq P\left\{(c - a)^2 L' e^{\lambda D} > e^{\lambda(D+\delta)}\right\}$$
$$= P\left\{L' > e^{\lambda\delta}/(c - a)^2\right\} \to 0, \quad \lambda \to \infty.$$

To obtain an estimate from below first we notice that

(1.3) $$\lim_{\lambda\to\infty} \lambda^{-1}\log C_\lambda = D,$$

3) $\max_I W = \max\{W(x), x \in I\}$, $\min_I W = \min\{W(x), x \in I\}$.

where

$$C_\lambda = |\hat{S}_\lambda(a)| \wedge |\hat{S}_\lambda(c)| ,$$

and the convergence is uniform in $x \in [u, v]$. Next, for given $\delta > 0$ we set

$$a_1 = \sup\{x < b_1 : W(x) = \delta/4\} ,$$
$$\hat{s}_\lambda(y) = \hat{S}_\lambda(y)/C_\lambda ,$$
$$L_\lambda = \min\{L(-1, 1, y) : \hat{s}_\lambda(a_1) \le y \le \hat{s}_\lambda(b_1)\} .$$

Then applying Lemma 1 and 2 we have

$$T_\lambda^x \overset{d}{=} C_\lambda \int_a^c L(\hat{s}_\lambda(a), \hat{s}_\lambda(c), \hat{s}_\lambda(y)) e^{-\lambda W(y)} dy$$

$$\ge C_\lambda \int_{a_1}^{b_1} L(-1, 1, \hat{s}_\lambda(y)) e^{-\lambda W(y)} dy$$

$$\ge e^{\lambda(D-\frac{\delta}{4})} (b_1 - a_1) L_\lambda \exp\left\{-\lambda \max_{[a_1, b_1]} W\right\}$$

$$= (b_1 - a_1) L_\lambda e^{\lambda(D-\frac{\delta}{2})} .$$

Since $\lambda^{-1}\log|\hat{s}_\lambda(a_1)|$ and $\lambda^{-1}\log|\hat{s}_\lambda(b_1)|$ converges to $\max_{[x\wedge a_1, x\vee a_1]} W - D$, $\max_{[x\wedge b_1, x\vee b_1]} W - D$, respectively, which are both negative, we have

$$\lim_{\lambda\to\infty} \hat{s}_\lambda(a_1) = \lim_{\lambda\to\infty} \hat{s}_\lambda(b_1) = 0 ,$$

the convergence being uniform in $x \in [u, v]$. Therefore

$$P\left\{T_\lambda^x < e^{\lambda(D-\delta)}\right\} \le P\left\{L_\lambda < (b_1 - a_1)^{-1} e^{-\lambda\delta/2}\right\} \to 0 , \quad \lambda \to \infty$$

uniformly in $x \in [u, v]$, because $\lim_{\lambda\to\infty} L_\lambda = L(-1, 1, 0) > 0$.

§2. The limit distribution of $X(e^{\lambda r}, \lambda W)$

In this section we change the notation slightly. Given $W \in \mathbb{W}$ and a valley $V = (a, b_1, b_2, c)$ of W , we set

$$\Omega = C([0, \infty) \to \mathbb{R}) ,$$
$$\hat{\Omega} = C([0, \infty) \to [a, c]) ,$$

and denote by P_λ^x , $x \in \mathbb{R}$ (resp. $\hat{P}_\lambda^y$, $y \in [a, c]$) the probability measure on Ω (resp. $\hat{\Omega}$) induced by the diffusion process with generator

$$\frac{1}{2} e^{\lambda W(x)} \frac{d}{dx}\left(e^{-\lambda W(x)} \frac{d}{dx}\right) \tag{2.1}$$

(resp. the diffusion process on $[a, c]$ with (local) generator (2.1) and with reflecting barriers at a and c). The latter diffusion has the invariant probability measure m_λ given by

$$m_\lambda(dy) = e^{-\lambda W(y)} dy \Big/ \int_a^c e^{-\lambda W(z)} dz .$$

For any interval $[u, v] \subset [a, c]$

$$m_\lambda([u, v]) = \frac{\int_0^\infty e^{-\lambda \xi} K([u, v], \xi) d\xi}{\int_0^\infty e^{-\lambda \xi} K([a, c], \xi) d\xi}$$

where, for an interval I in $\mathbb{R}$, $K(I, \xi)$ is the local time at ξ for the reflected Brownian medium, i.e.,

$$K(I, \xi) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int_I \mathbb{1}_{[\xi, \xi+\varepsilon)}(W(s)) ds . \tag{2.2}$$

Therefore

$$m_\lambda([u, v]) = \frac{\int_0^\infty e^{-\xi} K([u, v], \lambda^{-1}\xi) d\xi}{\int_0^\infty e^{-\xi} K([a, c], \lambda^{-1}\xi) d\xi} \tag{2.3}$$

$$\to \frac{K([u, v], 0)}{K([a, c], 0)} \equiv m([u, v]) , \quad \lambda \to \infty .$$

Next we set

$$\hat{P}_\lambda = \int_a^b m_\lambda(dy) \hat{P}^y_\lambda , \quad \mathbb{P}^{x,y}_\lambda = P^x_\lambda \otimes \hat{P}^y_\lambda , \quad \mathbb{P}^x_\lambda = P^x_\lambda \otimes \hat{P}_\lambda .$$

$$R = R(\omega, \hat{\omega}) = \inf\{t \geq 0 : \omega(t) = \hat{\omega}(t)\} .$$

<u>Lemma 5</u>. For any $\delta > 0$

$$\lim_{\lambda \to \infty} \mathbb{P}^0_\lambda\left\{R < e^{\lambda(A+\delta)}\right\} = 1 .$$

Proof. First we prove that

$$\lim_{\lambda \to \infty} \mathbb{P}^x_\lambda\left\{R < e^{\lambda(A+\delta)}\right\} = 1 \text{ holds for } x = b_1 \text{ and } b_2 . \tag{2.4}$$

Without loss of generality we may consider the case $x = b_2$. We write b instead of b_2 for simplicity. For any $\delta > 0$ such that $A + \delta < D$ we define $a_1 \in (a, b_1)$, $a_2 \in (a, b_1)$, $c_2 \in (b_2, c)$ by

$$a_1 = \max\left\{x < b_1 : W(x) = A + \frac{\delta}{4}\right\},$$
$$a_2 = \max\left\{x < b_1 : W(x) = A + \frac{\delta}{2}\right\},$$
$$c_2 = \min\left\{x > b_2 : W(x) = A + \frac{\delta}{2}\right\},$$

and set

$$T_0 = T_0(\omega) = \inf\left\{t \geq 0 : w(t) = a_1\right\},$$
$$T_1 = T_1(\omega) = \inf\left\{t \geq 0 : w(t) \notin (a_1, c_2)\right\},$$
$$T_2 = T_2(\omega) = \inf\left\{t \geq 0 : w(t) \notin (a_2, c_2)\right\}.$$

Then we can prove easily that

$$(2.5) \qquad P_\lambda^b\left\{T_0 < \infty\right\} \geq P_\lambda^b\left\{T_0 = T_1\right\} = \frac{S_\lambda(c_2) - S_\lambda(b)}{S_\lambda(c_2) - S_\lambda(a_1)} \to 1, \ \lambda \to \infty,$$

and hence

$$(2.6) \qquad \begin{aligned} &\mathbb{P}_\lambda^b\left\{R \leq T_0\right\} \\ &\geq \mathbb{P}_\lambda^b\left\{\hat{\omega}(0) \in [a, b], \hat{\omega}(T_0) \in [a_1, c]\right\} \\ &\geq \mathbb{P}_\lambda^b\left\{\hat{\omega}(0) \in [a, b]\right\} + P_\lambda^b\left\{\hat{\omega}(T_0) \in [a_1, c]\right\} - 1 \\ &= m_\lambda([a, b]) + \int_0^\infty \hat{P}_\lambda\left\{\hat{\omega}(t) \in [a_1, c]\right\} P_\lambda^b\left\{T_0 \in dt\right\} - 1 \\ &\to 1, \ \lambda \to \infty, \end{aligned}$$

by (2.3) because $m(\{x \in (a, c) : W(x) = 0\}) = 1$. On the other hand Lemma 4 applied to the valley (a_2, b_1, b_2, c_2) whose depth is $A + (\delta/2)$ implies

$$(2.7) \qquad P_\lambda^b\left\{T_1 < e^{\lambda(A+\delta)}\right\} \geq P_\lambda^b\left\{T_2 < e^{\lambda(A+\delta)}\right\} \to 1, \ \lambda \to \infty,$$

and so

$$\begin{aligned} &\mathbb{P}_\lambda^x\left\{R < e^{\lambda(A+\delta)}\right\} \\ &\geq P_\lambda^x\left\{T_0 < e^{\lambda(A+\delta)}\right\} - o(1) \qquad \text{(by (2.6))} \\ &\geq P_\lambda^x\left\{T_1 < e^{\lambda(A+\delta)}, \ T_1 = T_0\right\} - o(1) \end{aligned}$$

$$\geq P_\lambda^x\left\{T_1 < e^{\lambda(A+\delta)}\right\} - o(1) \qquad \text{(by (2.5))}$$

$$\longrightarrow 1 , \quad \text{as } \lambda \to \infty \qquad \text{(by (2.7))} .$$

Next , to consider the case where the diffusion starts at 0 we shall consider three diffusion processes starting at 0 , b_1 and b_2 , respectively. By making use of the comparison theorem in one-dimensional diffusion processes (for example, see [3: p.352]) we can construct, on a suitable probability space $(\widetilde{\Omega}_\lambda, \widetilde{P}_\lambda)$, three processes $\widetilde{X}_0(t)$, $\widetilde{X}_1(t)$ and $\widetilde{X}_2(t)$ such that the probability measure on Ω induced by $\widetilde{X}_0(t)$ (resp. $\widetilde{X}_1(t)$, $\widetilde{X}_2(t)$) coincides with P_λ^0 (resp. $P_\lambda^{b_1}$, $P_\lambda^{b_2}$) and

$$\text{(2.8)} \qquad \widetilde{X}_1(t) \leq \widetilde{X}_0(t) \leq \widetilde{X}_2(t) , \qquad \forall t \geq 0 , \ \widetilde{P}_\lambda\text{-a.s.}$$

Put

$$\widetilde{\mathbb{P}}_\lambda = \widetilde{P}_\lambda \otimes \widehat{P}_\lambda ,$$

$$\widetilde{R}_i = \inf\left\{t \geq 0 : \widetilde{X}_i(t) = \widehat{\omega}(t)\right\} , \quad i = 0, 1, 2 .$$

Since $\widetilde{R}_0 \leq \widetilde{R}_1 \vee \widetilde{R}_2$ by (2.8), we have

$$\begin{aligned} P_\lambda^0\left\{R < e^{\lambda(A+\delta)}\right\} &= \widetilde{\mathbb{P}}_\lambda\left\{\widetilde{R}_0 < e^{\lambda(A+\delta)}\right\} \\ &\geq \widetilde{\mathbb{P}}_\lambda\left\{\widetilde{R}_1 \vee \widetilde{R}_2 < e^{\lambda(A+\delta)}\right\} \\ &\geq \mathbb{P}_\lambda^{b_1}\left\{R < e^{\lambda(A+\delta)}\right\} + \mathbb{P}^{b_2}\left\{R < e^{\lambda(A+\delta)}\right\} - 1 \\ &\longrightarrow 1 , \quad \lambda \to \infty \end{aligned}$$

by (2.4), completing the proof of Lemma 5.

<u>Lemma 6</u>. For any r_1 , r_2 with $A < r_1 < r_2 < D$ and for any interval $[u, v]$ in $\mathbb{R}$

$$\lim_{\lambda \to \infty} P_\lambda^0\left\{\omega(e^{\lambda r}) \in [u, v]\right\} = m([u, v] \cap [b_1, b_2])$$

uniformly in $r \in [r_1, r_2]$, where m is defined in (2.3).

Proof. Denote by T (resp. $\widehat{T}$) the exit time of (a, c) for $\omega(t)$ (resp. $\widehat{\omega}(t)$), and by T_R (resp. $\widehat{T}_R$) the exit time of (a, c) for $\omega(t)$ (resp. $\widehat{\omega}(t)$) after the collision time R . Since $m_\lambda(U) \to 1$ as $\lambda \to \infty$ for any open set U containing $\{x \in (a, c) : W(x) = 0\}$, it follows from Lemma 4 that

$$\widehat{P}_\lambda\left\{e^{\lambda(D-\delta)} < \widehat{T} < e^{\lambda(D+\delta)}\right\}$$

$$= \int_a^c m_\lambda(dx) P_\lambda^x \left\{ e^{\lambda(D-\delta)} < T < e^{\lambda(D+\delta)} \right\}$$

$$\longrightarrow 1 \ , \ \lambda \to \infty .$$

This combined with Lemma 5 implies

$$p_\lambda \ := \mathbb{P}_\lambda^0 \left\{ R < e^{\lambda r_1} < e^{\lambda r_2} < \hat{T}_R \right\}$$

$$\geq \mathbb{P}_\lambda^0 \left\{ R < e^{\lambda r_1} < e^{\lambda r_2} < \hat{T} \right\} \quad (\because \hat{T} \leq \hat{T}_R)$$

$$\longrightarrow 1 \ , \ \lambda \to \infty .$$

Therefore for $r \in [r_1, r_2]$

$$\text{(2.9)} \qquad P_\lambda^0 \left\{ \omega(e^{\lambda r}) \in [u, v] \right\}$$

$$\geq \mathbb{P}_\lambda^0 \left\{ R < e^{\lambda r_1}, \omega(e^{\lambda r}) \in [u, v], \ e^{\lambda r_2} < \tilde{T}_R \right\}$$

$$= \mathbb{P}_\lambda^0 \left\{ R < e^{\lambda r_1}, \hat{\omega}(e^{\lambda r}) \in [u, v], \ e^{\lambda r_2} < \hat{T}_R \right\}$$

$$\geq p_\lambda + m_\lambda([u, v]) - 1$$

$$\longrightarrow m([u \ , \ v] \cap [b_1, \ b_2]) \ , \lambda \to \infty \ ;$$

as for the above equality we used the strong Markov property. Similarly we have

$$\varliminf_{\lambda \to \infty} P_\lambda^0 \left\{ \omega(e^{\lambda r}) \in [u, \ v]^c \right\} \geq m([u, \ v]^c \cap [b_1, \ b_2]) \ ,$$

which combined with (2.9) implies

$$P_\lambda^0 \left\{ \omega(e^{\lambda r}) \in [u, \ v] \right\} \to m([u, \ v] \cap [b_1, \ b_2]) \ , \lambda \to \infty \ .$$

The uniform convergence in $r \in [r_1, r_2]$ is also clear.

§3. Proof of Theorem 1

Let $V = (a, b_1, b_2, c)$ be a valley of W such that $A < 1 < D$. Such a valley exists with Q-probability 1. In fact, b_1 and b_2 are taken as

b_1 = the smallest root of $W(x) = 0$ in $(a', 0)$

b_2 = the largest root of $W(x) = 0$ in $(0, c')$

where $a' = \sup\{x < 0 : W(x) = 1\}$ and $c' = \inf\{x > 0 : W(x) = 1\}$. The endpoints a and c can be chosen suitably so that $a < a'$, $c > c'$ and

$V = (a, b_1, b_2, c)$ is a valley with $A < 1 < D$. In what follows $V = (a, b_1, b_2, c)$ denotes such a valley of W. We denote by m_W the probability measure on $\mathbb{R}$ defined by

$$m_W([u, v]) = \frac{K([u', v'], 0)}{K([b_1, b_2], 0)} \tag{3.1}$$

where $[u', v'] = [u, v] \cap [b_1, b_2]$. Then, in the notation of §1 Lemma 6 reads as follows: For any interval I in $\mathbb{R}$ and for any family $\{r(\lambda), \lambda > 0\}$ satisfying $\lim_{\lambda\to\infty} r(\lambda) = 1$,

$$\lim_{\lambda\to\infty} P\left\{X(e^{\lambda r(\lambda)}, \lambda W) \in I\right\} = m_W(I) \tag{3.2}$$

for almost all W with respect to Q. Now we define $\mathbb{P} = P \otimes Q$ and $\mu = \int m_W Q(dW)$. Integrating both sides of (3.2) with respect to Q we have

$$\lim_{\lambda\to\infty} \mathbb{P}\left\{X(e^{\lambda r(\lambda)}, \lambda W) \in I\right\} = \mu(I). \tag{3.3}$$

Next, define W_λ as in Lemma 3. Then $\{W_\lambda(x), x \in \mathbb{R}\}$ is again a reflected Brownian medium. Therefore (3.3) yields

$$\lim_{\lambda\to\infty} \mathbb{P}\left\{X(e^{\lambda r(\lambda)}, \lambda W_\lambda) \in I\right\} = \mu(I). \tag{3.4}$$

We now apply the scaling relation (1.2) to (3.4); the result is

$$\lim_{\lambda\to\infty} \mathbb{P}\left\{\lambda^{-2} X(\lambda^4 e^{\lambda r(\lambda)}, W) \in I\right\} = \mu(I).$$

Taking $r(\lambda) = 1 - 4\lambda^{-1}\cdot\log\lambda$ in the above, we obtain

$$\lim_{\lambda\to\infty} \mathbb{P}\left\{\lambda^{-2} X(e^{\lambda}, W) \in I\right\} = \mu(I).$$

This completes the proof of Theorem 1.

§4. Proof of Theorem 2

The absolute continuity of μ can be proved easily. In fact, if μ_n is the measure in $\mathbb{R}$ defined by

$$\mu_n(I) = E^Q\left\{\frac{K(I \cap [b_1, b_2])}{K([b_1, b_2])} ; K([b_1, b_2]) > \frac{1}{n}\right\},$$

then μ_n is absolutely continuous because

$$\begin{aligned}\mu_n(I) &\le n E^Q\left\{K(I \cap [b_1, b_2])\right\} \\ &= 2n \int_I p(|x|, 0, 0)\,dx,\end{aligned}$$

where $p(t, \xi, \eta)$ is the transition density of the Brownian motion with absorbing barriers at ± 1. Thus μ is absolutely continuous because $\mu_n \uparrow \mu$ as $n \uparrow \infty$.

We proceed to the proof of (3). Let $K(I) = K(I, 0)$ be the local time at 0 for the reflected Brownian medium as defined by (2.2) with $\xi = 0$ and consider the number of times $d_\varepsilon(t)$ that the reflected Brownian path $\{W(u) : u \geq 0\}$ crosses down from $\varepsilon > 0$ to 0 before time t. Then as found in [4: p.48]

$$Q\left\{\lim_{\varepsilon \downarrow 0} 2\varepsilon d_\varepsilon(t) = K([0, t]),\ t \geq 0\right\} = 1 . \tag{4.1}$$

Let a', c', b_1 and b_2 be defined as in the beginning of §3.

Lemma 7. For $\alpha, \beta > 0$

$$E^Q\left\{e^{-\alpha K([0, b_2]) - \beta c'}\right\} = \frac{1}{2\alpha + c(\beta)} \cdot \frac{2\sqrt{2\beta}}{e^{\sqrt{2\beta}} - e^{-\sqrt{2\beta}}} , \tag{4.2}$$

where

$$c(\beta) = \frac{e^{\sqrt{2\beta}} + e^{-\sqrt{2\beta}}}{e^{\sqrt{2\beta}} - e^{-\sqrt{2\beta}}} \cdot \sqrt{2\beta} .$$

In Particular, $K([0, b_2])$ is exponentially distributed:

$$E^Q\left\{e^{-\alpha K([0, b_2])}\right\} = \frac{1}{2\alpha + 1} . \tag{4.3}$$

Proof. Since $c(\beta) \sim 1$ as $\beta \downarrow 0$, (4.3) follows from (4.2) by letting $\beta \downarrow 0$. To prove (4.2) we first apply (4.1) to write down

$$\begin{aligned} &E^Q\left\{e^{-\alpha K([0, b_2]) - \beta c'}\right\} \\ &= E^Q\left\{e^{-\alpha K([0, c']) - \beta c'}\right\} \\ &= \lim_{\varepsilon \downarrow 0} E^Q\left\{e^{-2\alpha \varepsilon d_\varepsilon(c') - \beta c'}\right\} \\ &= \lim_{\varepsilon \downarrow 0} \sum_{n=0}^{\infty} e^{-2\alpha\varepsilon n}\, E^Q\left\{e^{-\beta T_\varepsilon}\right\}^{n+1} E^Q_\varepsilon\left\{e^{-\beta T_0} ; T_0 < T_1\right\}^n E^Q_\varepsilon\left\{e^{-\beta T_1} ; T_1 < T_0\right\} , \end{aligned} \tag{4.4}$$

where E^Q_ε denotes the expectation with respect to the probability measure of the reflected Brownian motion starting at ε and

$$T_x = \inf\left\{u \geq 0 : W(u) = x\right\} .$$

If we set

$$A_\varepsilon = e^{-2\alpha\varepsilon} E^Q\left\{e^{-\beta T_\varepsilon}\right\} E_\varepsilon^Q\left\{e^{-\beta T_0};\ T_0 < T_1\right\},$$

$$B_\varepsilon = E^Q\left\{e^{-\beta T_\varepsilon}\right\} E_\varepsilon^Q\left\{e^{-\beta T_1};\ T_1 < T_0\right\},$$

then (4.4) yields

$$(4.5)\qquad E^Q\left\{e^{-\alpha K([0,b_2])-\beta c'}\right\} = \lim_{\varepsilon\downarrow 0} B_\varepsilon \sum_{n=0}^{\infty} A_\varepsilon^n = \lim_{\varepsilon\downarrow 0} \frac{B_\varepsilon}{1 - A_\varepsilon}.$$

Next we make use of the well-known formula

$$E_x\left\{e^{-\alpha T_a}; T_a < T_b\right\} = \frac{e^{\sqrt{2\alpha}(b-x)} - e^{-\sqrt{2\alpha}(b-x)}}{e^{\sqrt{2\alpha}(b-a)} - e^{-\sqrt{2\alpha}(b-a)}}, \quad a \le x \le b,$$

where E_x denotes the expectation with respect to the probability measure of a standard Brownian motion starting at x. We then have

$$(4.6)\qquad E^Q\left\{e^{-\beta T_\varepsilon}\right\} = 2E_0\left\{e^{-\beta T_\varepsilon};\ T_\varepsilon < T_{-\varepsilon}\right\} = \frac{2(e^{\varepsilon\sqrt{2\beta}} - e^{-\varepsilon\sqrt{2\beta}})}{e^{2\varepsilon\sqrt{2\beta}} - e^{-2\varepsilon\sqrt{2\beta}}} = 1 + 0(\varepsilon^2), \quad \varepsilon\downarrow 0;$$

$$(4.7)\qquad E_\varepsilon^Q\left\{e^{-\beta T_0}; T_0 < T_1\right\} = \frac{e^{\sqrt{2\beta}(1-\varepsilon)} - e^{-\sqrt{2\beta}(1-\varepsilon)}}{e^{\sqrt{2\beta}} - e^{-\sqrt{2\beta}}} = 1 - \frac{\sqrt{2\beta}(e^{\sqrt{2\beta}} + e^{-\sqrt{2\beta}})}{e^{\sqrt{2\beta}} - e^{-\sqrt{2\beta}}}\cdot\varepsilon + o(\varepsilon), \quad \varepsilon\downarrow 0;$$

$$(4.8)\qquad E_\varepsilon^Q\left\{e^{-\beta T_1};\ T_1 < T_0\right\} = \frac{e^{\sqrt{2\beta}\,\varepsilon} - e^{-\sqrt{2\beta}\,\varepsilon}}{e^{\sqrt{2\beta}} - e^{-\sqrt{2\beta}}} \sim \frac{2\sqrt{2\beta}}{e^{\sqrt{2\beta}} - e^{-\sqrt{2\beta}}}\cdot\varepsilon, \quad \varepsilon\downarrow 0.$$

From (4.6), (4.7) and (4.8) we obtain

$$\frac{B_\varepsilon}{1 - A_\varepsilon} \sim \frac{1}{2\alpha + c(\beta)} \cdot \frac{2\sqrt{2\beta}}{e^{\sqrt{2\beta}} - e^{-\sqrt{2\beta}}} \quad , \ \varepsilon \downarrow 0 ,$$

which combined with (4.5) proves the lemma.

Given $x > 0$ we set

$$K_1 = K([b_1, 0]), \quad K_2 = K([0,x]), \quad K_3 = K([x, b_2]) .$$

Lemma 8. For $x > 0$ and $t > 0$

$$(4.9) \qquad E^Q\left\{K_3 e^{-t(K_1+K_2+K_3)} ; \ x < b_2\right\} = \frac{2}{(2t+1)^3} E^Q\left\{(1-W(x))e^{-tK([0,x])} ; \ x < c'\right\} .$$

Proof. The left hand side of (4.9) equals

$$E^Q\left\{e^{-tK_1}\right\} E^Q\left\{K_3 e^{-t(K_2+K_3)} ; \ x < b_2\right\} .$$

Since $E^Q\left\{e^{-tK_1}\right\} = (2t+1)^{-1}$ by Lemma 7, for the proof of the lemma it is enough to show

$$(4.10) \qquad E^Q\left\{K_3 e^{-t(K_2+K_3)} : \ x < b_2\right\} = \frac{2}{(2t+1)^2} E^Q\left\{(1 - W(x))e^{-tK_2} ; \ x < c'\right\} .$$

To prove this we introduce the smallest σ-field $\mathcal{F}_x$ on $\mathbb{W}$ which makes $W(s)$, $0 \le s \le x$, measurable and consider the event Γ that the shifted trajectory $W(\cdot + x)$ hits 0 before it hits 1. Then first using the strong Markov property of the reflected Brownian motion and then (4.3), we have

$$E^Q\left\{K_3 e^{-tK_3} \mathbb{1}_\Gamma / \mathcal{F}_x\right\} = \left\{1 - W(x)\right\} E^Q\left\{K([0, b_2]) e^{-tK([0,b_2])}\right\} = \frac{2}{(2t+1)^2}\left\{1 - W(x)\right\} , \quad \text{a.s.}$$

Since $\left\{x < b_2\right\} = \left\{x < c'\right\} \cap \Gamma$ and $\left\{x < c'\right\} \in \mathcal{F}_x$, we have

$$E^Q\left\{K_3 e^{-t(K_2+K_3)} ; \ x < b_2\right\} = E^Q\left(e^{-tK_2} \mathbb{1}_{\{x<c'\}} E^Q\left\{K_3 e^{-tK_3} \mathbb{1}_\Gamma / \mathcal{F}_x\right\}\right)$$

$$= \frac{2}{(2t+1)^2} E^Q\left\{(1 - W(x))e^{-tK_2}; \ x < c'\right\} ,$$

proving (4.10) and hence the lemma.

Lemma 9. For $\lambda > 0$ and $t > 0$

$$(4.11) \qquad \int_0^\infty e^{-\lambda x} E^Q\left\{(1 - W(x))e^{-tK([0,x])}; \ x < c'\right\} dx$$
$$= \frac{1}{\lambda}\left\{1 - \frac{(2t+1)S}{C + 2tS}\right\},$$

where

$$C = \cosh\sqrt{2\lambda} \quad , \quad S = \frac{\sinh\sqrt{2\lambda}}{\sqrt{2\lambda}} .$$

Proof. Let $\varphi(x) = 1 - |x|$. Consulting with [4: Chapter 5] , we see that the left hand side of (4.11) equals $f_\lambda(0)$ where f_λ is the continuous solution of

$$(4.12) \qquad \begin{cases} \lambda f - \frac{1}{2}f'' = \varphi & \text{in } (-1, 0)\cup(0, 1) \\ \frac{1}{2}\{f'(0+) - f'(0-)\} = 2tf(0) \\ f(-1) = f(1) = 0 . \end{cases}$$

To solve (4.12) we first find the solution g_λ of $\lambda f - \frac{1}{2}f'' = \varphi$ in $(-1,1)$ with boundary condition $f(-1) = f(1) = 0$ and then express f_λ as follows:

$$f_\lambda(x) = \begin{cases} g_\lambda(x) + c \sinh\sqrt{2\lambda}(1+x) & \text{for } x \in (-1, 0) \\ g_\lambda(x) + c \sinh\sqrt{2\lambda}(1-x) & \text{for } x \in (0, 1) \end{cases} .$$

If we determine c so that the above f_λ satisfies the second condition of (4.12), then the f_λ is a solution of (4.12). Thus $f_\lambda(0)$ can be computed and we obtain (4.11).

Now Theorem 2 can be proved as follows. By Lemma 8 we have

$$\mu((x,\infty)) = E^Q\left\{\frac{K((x, x\vee b_2])}{K([b_1, b_2])}\right\}$$
$$= E^Q\left\{\frac{K_3}{K_1 + K_2 + K_3} ; \ x < b_2\right\}$$
$$= \int_0^\infty E^Q\left\{K_3 e^{-t(K_1+K_2+K_3)}; \ x < b_2\right\} dt$$

$$= \int_0^\infty \frac{2}{(2t+1)^3} E^Q\left\{(1 - W(x))e^{-tK([0,x])} \ ; \ x < c'\right\}dt$$

and hence by Lemma 9

$$\int_0^\infty e^{-\lambda x}\mu((x,\infty))dx = \int_0^\infty \frac{2}{(2t+1)^3}\cdot\frac{1}{\lambda}\left\{1 - \frac{(2t+1)S}{C+2tS}\right\}dt$$

$$= \frac{1}{2\lambda} - \frac{1}{\lambda}\int_0^\infty \frac{2}{(2t+1)^2}\cdot\frac{S}{C+2tS}\,dt \ .$$

Thus integration by parts yields

$$\int_0^\infty e^{-\lambda x}\mu(dx) = \frac{1}{2} - \lambda\int_0^\infty e^{-\lambda x}\mu((x,\infty))dx \quad \text{(notice that } \mu((0,\infty)) = \tfrac{1}{2})$$

$$= \int_0^\infty \frac{2S}{(2t+1)^2(C+2tS)}\,dt$$

$$= \int_0^\infty \frac{S\,dt}{(t+1)^2(C+tS)} \quad ,$$

and this proves (3).

REFERENCES

[1] T.Brox, A one-dimensional diffusion process in a Wiener medium, to appear in Ann. Probab.

[2] A.O.Golosov, The limit distributions for random walks in random environments, Soviet Math. Dokl., 28(1983), 18-22.

[3] N.Ikeda and S.Watanabe, Stochastic Differential Equations and Diffusion Processes, North-Holland, 1981.

[4] K.Itô and H.P.McKean, Diffusion Processes and Their Sample Paths, Springer-Verlag, 1965.

[5] K.Kawazu, Y.Tamura and H.Tanaka, One-dimensional diffusions and random walks in random environments, in preparation.

[6] H.Kesten, The limit distribution of Sinai's random walk in random environment, to appear in Physica.

[7] S.Schumacher, Diffusions with random coefficients, Contemporary Math. (Particle Systems, Random Media and Large Deviations, ed. by R. Durrett), 41(1985), 351-356.

[8] Y.G.Sinai, The limiting behavior of a one-dimensional random walk in a random medium, Theory of Probab. and its Appl. 27(1982), 256-268.
[9] H.Tanaka, Limit distributions for one-dimensional diffusion processes in self-similar random environments, to appear in the Proc. of the workshop on Hydrodynamic behavior and interacting particle systems and applications held at IMA, University of Minnesota, March 1986.

Department of Mathematics
Faculty of Science and Technology
Keio University
Yokohama, Japan

Reprinted from
Seminaire de Probabilites XXI, 246–261,
Lecture Notes in Math., **1247**, Springer-Verlag, 1987.

LIMIT DISTRIBUTIONS FOR ONE-DIMENSIONAL DIFFUSION PROCESSES IN SELF-SIMILAR RANDOM ENVIRONMENTS

H. Tanaka

Department of Mathematics
Faculty of Science and Technology
Keio University, Yokohama, JAPAN

Introduction

Let $X(t)$ be the one-dimensional diffusion process described by the stochastic differential equation

(1) $$dX(t) = dB(t) - \frac{1}{2} W'(X(t))dt, \qquad X(0) = 0,$$

where $B(t)$ is a one-dimensional Brownian motion starting at 0 and $\{W(x), x \in \mathbb{R}\}$ is a random environment which is independent of the Brownian motion $B(t)$. We are interested in the asymptotic behavior of $X(t)$ as $t \to \infty$: Under what scaling does $X(t)$ have a limit distribution? Similar problems for random walks were considered by Kesten, Kozlov and Spitzer [5] and Sinai [8]. The problem we discuss here is a diffusion analogue of Sinai's random walk problem [8]. In the case of a Brownian environment Brox [1] proved that the distribution of $(\log t)^{-2}X(t)$ is convergent as $t \to \infty$. Similar results were obtained by Schumacher [7] for a considerably wider class of self-similar random environments. As was seen by these works the assumption of the self-similarity of the random environment is important and the notion of suitably defined valleys of the environment plays a central role in the proof. However, it was assumed that the environment has only one point which gives the same value of local minima or maxima (the bottom of a valley consists of a single point), and the explicit form of the limit distribution was unknown until a recent discovery by Kesten ([6]) for Sinai's random walk which corresponds to the case of a Brownian environment in our diffusion setup (Golosov also obtained the same result as Kesten's; see also Golosov [2] for the corresponding result in another different model).

In this paper we discuss the following three typical examples of random environments with emphasis on finding the limit distributions:

(i) Nonpositive reflected Brownian environment.

(ii) Nonnegative reflected Brownian environment.

(iii) Symmetric stable environment.

In the first two examples the environment has (uncountably) many points giving the same value of local maxima or minima. The proof in (ii) is only sketched. For details see [9]. The result in (iii) on the limit distribution is an extension of Kesten's result [6].

In [4] a unified definition of valleys of an environment is given in a general setup containing the above examples and some results similar to Brox's and Schumacher's are obtained, but here we limit ourselves to the above examples because we are interested mainly in the form of the limit distributions and we have explicit results concerning this only in some special cases.

The author wishes to thank K. Kawazu and Y. Tamura: his frequent discussions with them were very valuable.

1. Preliminaries and the Result of Brox

1.1. Given a real valued right continuous function $W(x)$ defined on the real line $\mathbb{R}$ and having left limits, we consider the stochastic differential equation (1) with environment $W(x)$. Since $W(x)$ is not differentiable in general, what is meant by a solution of (1) will need explanation. However, without considering a solution itself for a given Brownian motion $B(t)$, we just *interprete the diffusion described by (1) as a diffusion process starting* at 0 with generator

$$\frac{1}{2} e^{W(x)} \frac{d}{dx} \left(e^{-W(x)} \frac{d}{dx}\right). \tag{1.1}$$

Such a diffusion can be obtained from a Brownian motion through a scale change and a time change ([3]). To be precise let

$$\Omega = C([0,\infty) \to \mathbb{R})^{1)} \cap \{\omega(0) = 0\},$$

$$P = \text{the Wiener measure on } \Omega,$$

$$B(t) = \omega(t) = \text{the value of } \omega \text{ at time } t,$$

$$L(t,x) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int_0^t 1_{[x,x+\varepsilon)}(B(s))ds \qquad \text{(local time)},$$

$$S(x) = \int_0^x e^{W(y)}dy ,$$

$$A(t) = \int_0^t e^{-2W(S^{-1}(B(s)))}ds = \int_{\mathbb{R}} e^{-2W(S^{-1}(x))}L(t,x)dx, \ t > 0,$$

S^{-1}, A^{-1} = the inverse functions.

Then the process $X(t,W) = S^{-1}(B(A^{-1}(t))$ defined on the probability space (Ω,P) is a diffusion process with generator (1.1) starting at 0. The Brownian motion $B(t)$ used here is not the same as the one in (1) but we use the same notation. Let $(W^x)(\cdot) = W(\cdot+x)$. For a fixed $x \in \mathbb{R}$ we replace W in $X(t,W)$ by W^x and then consider

$$X^x(t,W) = x + X(t,W^x).$$

Then $X^x(t,W)$ is a diffusion process with generator (1.1) starting at x. In this paper we shall deal with $X(t) = X(t,W)$ and $X^x(t) = X^x(t,W)$ and the following notation will be used throughout.

$$S_\lambda(x) = \int_0^x e^{\lambda W(u)}du,$$

W_λ: a function (environment) defined by $W_\lambda(x) = \lambda^{-1}W(\lambda^2 x)$ for $\forall x \in \mathbb{R}$, $\lambda > 0$ being fixed.

In the following lemma due to Brox the medium W is fixed. For the proof see [1].

Lemma 1.1 ([1]). For each $\lambda > 0$

$$\{X(t,\lambda W_\lambda),\ t \geq 0\} \stackrel{d}{=} \{\lambda^{-2}X(\lambda^4 t,\ W),\ t \geq 0\}$$

where $\stackrel{d}{=}$ means the equality in distribution.

1) For a topological space R the notation $C([0,\infty) \to R)$ stands for the space of R-valued continuous functions defined on $[0,\infty)$.

When we consider the environment to be random, we denote by Q the probability distribution (on the space of environments) of the random environment. Since we are assuming that the random environment and the Brownian motion $B(t)$ are independent, the full distribution is $\mathcal{P} = P \times Q$. Thus, when the environment is fixed $X(t)$ is governed by P, and when the environment is random $X(t)$ is governed by $\mathcal{P}$.

1.2. In this subsection we limit ourselves to the case of continuous environment and state a result of Brox [1] in a form which is convenient for our use in §2.

Given a continuous function W on $\mathbb{R}$ which is supposed to be the environment in (1), we define a valley of W following [1]. For $x \neq y$ we put

$$H_{x,y} = \sup \{W(y') - W(x')\} \tag{1.2}$$

where the supremum is taken over all pairs of x' and y' such that $x < x' < y' < y$ or $y < y' < x' < x$ according as $x < y$ or $y < x$. A triple $V = (a,b,c)$ is called a <u>valley</u> of W if

(i) $a < b < c$,

(ii) $W(b) < W(x) < W(a)$ for every $x \in (a,b)$,
$W(b) < W(x) < W(c)$ for every $x \in (b,c)$,

(iii) $H_{a,b} < W(c) - W(b)$, $H_{c,b} < W(a) - W(b)$.

For a valley $V = (a,b,c)$, $D = (W(a) - W(b)) \wedge (W(c) - W(b))$ and $A = H_{a,b} \vee H_{c,b}$ are called the <u>depth</u> of V and the <u>inner directed ascent</u> of V, where $u \wedge v$ (resp. $u \vee v$) denotes $\min \{u,v\}$ (resp. $\max \{u,v\}$). It is obvious that $A < D$.

<u>Theorem 1.1 (Brox [1])</u>. (i) Let $V = (a,b,c)$ be a valley of W with the depth D and the inner directed ascent A.

(i) Let T_λ^x be the exit time of (a,c) for the process $X_\lambda^x(t) = x + X(t, \lambda W^x)$. Then for any $\delta > 0$ and any closed interval $I \subset (a,c)$

$$\lim_{\lambda \to \infty} \inf_{x \in I} P\{e^{\lambda(D-\delta)} < T_\lambda^x < e^{\lambda(D+\delta)}\} = 1.$$

(ii) For any $\varepsilon>0$, any closed interval $I\subset(a,c)$ and for any closed interval $J\subset(A,D)$

$$\lim_{\lambda\to\infty}\sup_{\substack{x\in I\\ r\in J}} P\{|X^x(e^{\lambda r}, \lambda W) - b| > \varepsilon\} = 0.$$

First proving the above result and then making use of the scaling relation (Lemma 1.1), Brox derived his main theorem:

Theorem 1.2 (Brox [1]). For any $\varepsilon>0$

$$P\{|\lambda^{-2}X(e^{-\lambda}, W) - b_\lambda| > \varepsilon\} \to 0, \ \lambda\to\infty$$

in probability with respect to the probability measure Q of the Brownian environment, where b_λ is the unique bottom of a valley $V_\lambda = (a_\lambda, b_\lambda, c_\lambda)$ of W_λ such that

(1.3) $$a_\lambda < 0 < c_\lambda \quad , \quad A_\lambda < 1 < D_\lambda.$$

2. Nonpositive Reflected Brownian Environment

As a simple example in which the environment has (uncountably) many points giving the same value of local maxima we consider the case of a nonpositive reflected Brownian environment. Main ideas of this section grew out from the conversation with Y. Tamura.

We first introduce the space $\underline{\underline{W}} = C(\mathbb{R}\to(-\infty, 0])\cap\{W(0) = 0\}$ and denote by Q the probability measure on $\underline{\underline{W}}$ with respect to which $\{W(u) : u\geq 0\}$ and $\{W(-u) : u\geq 0\}$ are independent reflected Brownian motions on $(-\infty,0]$. An essential difference between the present case and Brox's one is that there is no valley of W_λ satisfying (1.3) in the present case. For $W\in\underline{\underline{W}}$ we put

$$W^*(u) = W(u) - \inf_{[0,u]} W, \; u>0 ,$$
$$\qquad = W(u) - \inf_{[u,0]} W, \; u<0 ,$$

and define $a_\pm = a_\pm(W)$, $b_\pm = b_\pm(W)$ and $c_\pm = c_\pm(W)$ as follows.

1. $d_\pm = \pm \min \{u > 0 : W^*(\pm u) = 1\}$,

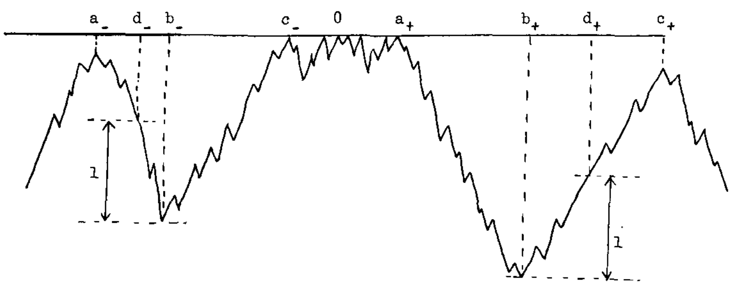
2. $b_\pm = \pm \min \{u > 0 : W(\pm u) = M_\pm\}$
   where $M_+ = \min_{[0,d_+]} W$ and $M_- = \min_{[d_-,0]} W$,

3. $c_- = \min \{u > b_- : W(u) = 0\}$,
   $a_+ = \max \{u < b_+ : W(u) = 0\}$,

4. $e_\pm = \pm \min \{u > d_\pm : W(\pm u) = W(b_\pm)\}$,

5. $a_- = \max \{u < b_- : W(u) = \max_{[e_-,b_-]} W\}$,
   $c_+ = \min \{u > b_+ : W(u) = \max_{[b_+,e_+]} W\}$.

Then $V_\pm = (a_\pm , b_\pm , c_\pm)$ are valleys of W with $a_- < b_- < c_- < 0 < a_+ < b_+ < c_+$, Q-a.s. As for the depth and the inner directed ascent we have $A_\pm < 1 < D_\pm$, Q-a.s., because $W(a_-) > W(d_-)$ and $W(c_+) > W(d_+)$, Q-a.s. Substracting a suitable null set from $\underline{\underline{W}}$ we may assume that the statements hold for all W in the subtracted space and so we often omit the phrase "Q-a.s."



Our aim is to know the asymptotic behavior of $X(t,W)$ as $t\to\infty$, or more precisely, to find the limit distribution of $(\log t)^{-2}X(t,W)$ as $t\to\infty$. First we observe the process $X(t,\lambda W)$. If λ becomes large and time goes on, the process eventually falls into one of the valleys $V_\pm$ and thereafter Theorem 1.1 will tell about the asymptotic behavior of $X(e^{\lambda r}, \lambda W)$, $\lambda \to \infty$, $r \sim 1$. Then a use of the scaling relation (Lemma 1.1) will give the result for the limit distribution of $\lambda^{-2}X(e^\lambda,W)$ as $\lambda\to\infty$.

To be precise choose α such that

$$(2.1) \qquad -\min_{[c_-,a_+]} W < \alpha < 1$$

and put

$$x_\pm = \pm \min\{u > 0:\ W(\pm u) = -\alpha\},$$
$$T_\lambda^\pm = \min\{t > 0: X(t,\lambda W) = x_\pm\},\ \hat{T}_\lambda = T_\lambda^- \quad T_\lambda^+ .$$

Denote by $L_-(I,\xi)$ the local time at ξ for the reflected Brownian environment $\{W(u),\ u\in\mathbb{R}\}$, i.e.,

$$L_-(I,\xi) = \lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\int_I 1_{(\xi-\varepsilon,\xi]}(W(u))du,\ I = \text{interval in } \mathbb{R}.$$

Lemma 2.1. (i) There exists $\delta = \delta(W) > 0$ such that

$$\lim_{\lambda\to\infty} P\{\hat{T}_\lambda < e^{\lambda(1-\delta)}\} = 1.$$

$$\text{(ii)} \qquad \lim_{\lambda\to\infty} P\{T_\lambda^- < T_\lambda^+\} = L_-([0,a_+],0)/L_-([c_-,a_+],0).$$

Proof. The condition (2.1) implies the existence of a modification $W^\#$ of W with the following properties (2.2) and (2.3).

(2.2) $\quad W^\#\in C(\mathbb{R}\to\mathbb{R})$ and $W^\#(x) = W(x)$ for $\forall x\in[x_-,x_+]$.

(2.3) $\quad$ There exists a valley $V^\# = (a^\#, b^\#, c^\#)$ of $W^\#$ with depth $D^\# < 1$ and $a^\# < x_- < x_+ < c^\#$.

Let $T_\lambda^\# = \min\{t > 0: X(t,\lambda W^\#)\notin (x_-, x_+)\}$. Then for any $\delta>0$ such that $D^\# + 2\delta < 1$ we have

$$P\{\hat{T}_\lambda < e^{\lambda(1-\delta)}\} = P\{T^{\#}_\lambda < e^{\lambda(1-\delta)}\} \geq P\{T^{\#}_\lambda < e^{\lambda(D^{\#}+\delta)}\} \to 1, \quad \lambda\to\infty$$

by (i) of Theorem 1.1. The proof of (ii) is easy; in fact

$$P\{T^-_\lambda < T^+_\lambda\} = \frac{S_\lambda(x_+) - S_\lambda(0)}{S_\lambda(x_+) - S_\lambda(x_-)}$$

$$= \frac{\int_{-\infty}^{0} e^{\lambda\xi} L_-([0, x_-], \xi)d\xi}{\int_{-\infty}^{0} e^{\lambda\xi} L_-([x_-, x_+], \xi)d\xi}$$

$$= \frac{\int_{-\infty}^{0} e^{\xi} L_-([0,x_+], \lambda^{-1}\xi)d\xi}{\int_{-\infty}^{0} e^{\xi} L_-([x_-,x_+], \lambda^{-1}\xi)d\xi}$$

$$\to L_-([0,x_+], 0)/L_-([x_-,x_+], 0), \quad \lambda \to \infty$$

$$= L_-([0,a_+], 0)/L_-([c_-, a_+], 0).$$

Lemma 2.2. There exist r_1 and r_2 with $r_1 < 1 < r_2$ such that for any $\varepsilon > 0$ and $r\in[r_1,r_2]$

$$\lim_{\lambda\to\infty} P\{X(e^{\lambda r}, \lambda W)\in U_\varepsilon(b)\} = p$$

holds with $b = b_\pm$ and $p = p_\pm$ where $p_+ = L_-([c_-,0],0)/L_-([c_-,a_+],0)$, $p_- = 1 - p_+$ and $U_\varepsilon(b)$ denotes the ε-neighborhood of b. The above convergence is uniform with respect to r on $[r_1, r_2]$.

Proof. It is easy to see that there exist valleys $\tilde{V}_\pm = (\tilde{a}_\pm, \tilde{b}_\pm, \tilde{c}_\pm)$ of W with depth $\tilde{D}_\pm > 1$ and satisfying $a_- < \tilde{a}_- < b_- < x_- < \tilde{c}_- < c_-$ and $a_+ < \tilde{a}_+ < x_+ < b_+ < \tilde{c}_+ < c_+$, respectively. Denote by $\tilde{T}^{x+}_\lambda$ the exit time of $\tilde{I}_+ = (\tilde{a}_+, \tilde{c}_+)$ for the process $X^{x_+}(t, \lambda W)$ and by $\tilde{T}^{x-}_\lambda$ the exit time of $\tilde{I}_- = (\tilde{a}_-, \tilde{c}_-)$ for the process $X^{x_-}(t,\lambda W)$. Then using the strong Markov property of $X(t,\lambda W)$ we have

$$(2.2) \qquad \begin{aligned} &P\{X(e^{\lambda(1-\delta)}, \lambda W)\in \tilde{I}_+\} \\ &\geq P\{T^+_\lambda = \hat{T}_\lambda < e^{\lambda(1-\delta)}\}\, P\{\tilde{T}^{x+}_\lambda > e^{\lambda(1-\delta)}\} \end{aligned}$$

and a similar inequality with + replaced by - . We now take $\delta = \delta(W)$ of Lemma 2.1 and then let $\lambda\uparrow\infty$ in (2.2). Then the assertion (i) of Theorem 1.1 combined with Lemma 2.1 implies

(2.3a) $$\varliminf_{\lambda\to\infty} P\{X(e^{\lambda(1-\delta)}, \lambda W)\in \Upsilon_+\} \geq p_+$$

and a similar inequality (= (2.3b)) with + replaced by - . (2.3a) and (2.3b) clearly imply that

(2.4) $$\lim_{\lambda\to\infty} P\{X(e^{\lambda(1-\delta)}, \lambda W)\in \Upsilon_\pm\} = p_\pm .$$

We now take r_1 and r_2 such that

$$A_-\vee A_+\vee(1-\delta) < r_1 < 1 < r_2 < D_-\wedge D_+$$

and prove the lemma with these r_1 and r_2. Let $r_1 < r < r_2$ and $t = e^{\lambda r} - e^{\lambda(1-\delta)}$. Then we have

$$P\{X(e^{\lambda r}, \lambda W)\in U_\varepsilon(b_+)\}$$
$$= \int_{\Upsilon_+} P\{X(e^{\lambda(1-\delta)}, \lambda W)\in dx\}P\{X^x(t,\lambda W)\in U_\varepsilon(b_+)\} + o(1)$$

which tends to p_+ uniformly in r as $\lambda\to\infty$ by Theorem 1.1 as applied to the valley V_+ because t can be expressed as $t = e^{\lambda r'}$ with $r'\to r$ as $\lambda\to\infty$. The other case ($b = b_-$, $p = p_-$) can be proved similarly and so the proof is finished.

Since $b_\pm$ and $p_\pm$ are Borel functions of W we can define a probability measure μ_- on R by

$$\int\phi(x)\mu_-(dx) = \int\{p_-\,\phi(b_-) + p_+\phi(b_+)\}Q(dW), \quad \phi\in C_b(\mathbb{R}).$$

<u>Theorem 2.1</u>. The full distribution of $(\log t)^{-2}X(t)$ converges to μ_- as $t\to\infty$.

Proof. Lemma 2.2 implies that the full distribution of $X(e^{\lambda r(\lambda)}, \lambda W)$ converges to μ_- as $\lambda\to\infty$ provided that $r(\lambda)\to 1$ as $\lambda\to\infty$. Since our reflected Brownian environment is self-similar, i.e., $\{W_\lambda(x), x\in\mathbb{R}\}$ is again a reflected

Brownian environment, we see that the full distribution of $X(e^{\lambda r(\lambda)}, \lambda W_\lambda)$ also converges to μ_-. Applying the scaling relation (Lemma 1.1) we see that the full distribution of $\lambda^{-2}X(\lambda^4 e^{\lambda r(\lambda)}, W)$ converges to μ_-. If we take $r(\lambda) = 1 - 4\lambda^{-1}\log\lambda$, the last statement is nothing but the assertion of the theorem.

Theorem 2.2. μ_- has a symmetric density and

$$(2.5) \qquad \int_0^\infty e^{-\lambda x}\mu_-(dx) = \int_0^\infty \frac{d\sigma}{(\sigma+1)^2(\sigma + \sqrt{2\lambda}\coth\sqrt{2\lambda})\cosh\sqrt{2\lambda}}, \quad \lambda > 0.$$

Proof. If $d_\varepsilon(t)$ denotes the number of times that the reflected Brownian path $\{W(u) : u \geq 0\}$ crosses down from 0 to $-\varepsilon$ before t, then $\lim_{\varepsilon \downarrow 0} 2\varepsilon d_\varepsilon(t) = L_-([0,t],0)$, $t \geq 0$, Q-a.s. (see [3]). Therefore, putting $f_+ = \min\{u>0 : W(u) = -1\}$ and $L_2 = L([0,a_+],0)$ we can write for $\sigma, \lambda > 0$

$$(2.6) \qquad E^Q\{e^{-\sigma L_2 - \lambda f_+}\}$$

$$= \lim_{\varepsilon \downarrow 0} \sum_{n=0} e^{-\sigma(2\varepsilon n)} E_0^Q\{e^{-\lambda H_{-\varepsilon}}\}^n E_{-\varepsilon}^Q\{e^{-\lambda H_0}; H_0 < H_{-1}\}^{n-1}$$

$$\cdot E_{-\varepsilon}^Q\{e^{-\lambda H_{-1}}; H_{-1} < H_0\},$$

where H_a denotes the hitting time to a for the reflected Brownian motion on $(-\infty,0]$ and the suffix x in E_x^Q indicates that the initial position is x. Using the explicit form of $E^Q\{e^{-\lambda H_{-\varepsilon}}\}$, etc., the right hand side of (2.6) can be computed. The result is

$$(2.7) \qquad E^Q\{e^{-\sigma L_2 - \lambda f_+}\} = \frac{2\sqrt{2\lambda}}{e^{\sqrt{2\lambda}} - e^{-\sqrt{2\lambda}}} \cdot \frac{1}{2\sigma + c(\lambda)},$$

where $c(\lambda) = (e^{\sqrt{2\lambda}} + e^{-\sqrt{2\lambda}})(e^{\sqrt{2\lambda}} - e^{-\sqrt{2\lambda}})^{-1}\sqrt{2\lambda}$. In particular L_2 is exponentially distributed with mean 2. In a similar spirit we see easily that

$$(2.8) \qquad E^Q\{e^{-\lambda g_+}\} = \frac{e^{\sqrt{2\lambda}} - e^{-\sqrt{2\lambda}}}{\sqrt{2\lambda}(e^{\sqrt{2\lambda}} + e^{-\sqrt{2\lambda}})},$$

where $g_+ = b_+ - f_+$. Let $L_1 = L_-([c_-, 0], 0)$. Then

$$\int_0^\infty e^{-\lambda x}\ \mu_-(dx) = E^Q\{\frac{L_1}{L_1 + L_2} \cdot e^{-\lambda b_+}\}$$

$$= \int_0^\infty E^Q\ \{L_1 e^{-\sigma(L_1 + L_2) - \lambda f_+ - \lambda g_+}\} d\sigma\ ,$$

and making use of (2.7), (2.8) and the fact that L_1, g_+ and $\{L_2, f_+\}$ are independent we can compute the right hand side of the above. We thus obtain (2.5).

3. Nonnegative Reflected Brownian Environment

In this section we consider the case of a nonnegative reflected Brownian environment. This is a typical case where the bottom of a valley consists of (uncountably) many points.

Let $\underline{\underline{W}}$ be the space $C(\mathbb{R} \to [0,\infty)) \cap \{W(0) = 0\}$ and consider the probability measure Q on $\underline{\underline{W}}$ with respect to which $\{W(u):u \geq 0\}$ and $\{W(-u):u \geq 0\}$ are independent reflected Brownian motions on $[0,\infty)$. The study of asymptotic behaviors of $X(t,W)$ as $t \to \infty$ can be done by a method similar to that of Brox [1] as will be sketched here.

The definition of a valley given in 1.2 must be slightly modified. Given $W \in \underline{\underline{W}}$, a quartet $V = (a, b_1, b_2, c)$ is called a <u>valley</u> of W if

(i) $a < b_1 < 0 < b_2 < c$,

(ii) $W(b_1) = W(b_2) = 0$, $W(a) = W(b) = D > 0$,

(iii) $0 < W(x) < W(a)$ for $0 < x < b_1$,

$0 < W(x) < W(c)$ for $b_2 < x < c$,

(iv) $A = H_{a,b_2} \vee H_{c,b_1} < D$.

D and A are called the <u>depth</u> and the <u>inner directed ascent</u> of V, respectively. There exists a valley of W such that $A < 1 < D$ with Q-probability 1. In fact, let

$$a' = \max\{u < 0 : W(u) = 1\}\ ,\quad c' = \min\{u > 0 : W(u) = 1\}\ ,$$

$$b_1 = \min\{u > a' : W(u) = 0\}\ ,\quad b_2 = \max\{u < c' : W(u) = 0\}\ .$$

Then with a suitable choice of a and c with $a < a'$, $c' < c$, the quartet $V = (a, b_1, b_2, c)$ becomes a valley of W with $A < 1 < D$, Q-a.s. In what follows $V = (a, b_1, b_2, c)$ denotes such a valley of W. We first observe $X(t, \lambda W)$. As in [1] we can prove that $X(e^{\lambda r}, \lambda W)$ falls into an ε-neighborhood of $[b_1, b_2]$ as $\lambda \to \infty$, $r \sim 1$. Then how is $X(e^{\lambda r}, \lambda W)$ distributed on $[b_1, b_2]$ in the limit? This limit distribution can be identified with the limit, as $\lambda \to \infty$, of the invariant probability measure m_λ of the diffusion process on $[a,c]$ with (local) generator $\frac{1}{2} e^{\lambda W(x)} \frac{d}{dx} (e^{-\lambda W(x)} \frac{d}{dx})$ and with reflecting barriers at a and c. If $L_+(I, \xi)$ denotes the local time at ξ for the reflected Brownian environment $\{W(x), x \in \mathbb{R}\}$, then for an interval $I \subset [a,c]$

$$m_\lambda(I) = \int_I e^{-\lambda W(y)} dy \Big/ \int_a^c e^{-\lambda W(y)} dy$$

$$= \int_0^\infty e^{-\lambda \xi} L_+(I, \xi) d\xi \Big/ \int_0^\infty e^{-\lambda \xi} L_+([a,c], \xi) d\xi.$$

$$\to L_+(I \cap [b_1, b_2], 0) / L_+([b_1, b_2], 0).$$

Define a probability measure m_W in $\mathbb{R}$ by

$$m_W([u,v]) = L_+(I', 0) / L_+([b_1, b_2], 0)$$

where $I' = [u,v] \cap [b_1, b_2]$. Then for any interval I in $\mathbb{R}$ and for any family $\{r(\lambda), \lambda > 0\}$ satisfying $\lim_{\lambda \to \infty} r(\lambda) = 1$ we have

$$\lim_{\lambda \to \infty} P\{X(e^{\lambda r(\lambda)}, \lambda W) \in I\} = m_W(I)$$

for almost all W with respect to Q. Now we define $\mu_+ = \int m_W Q(dW)$. Then

$$\lim_{\lambda \to \infty} \mathcal{P}\{X(e^{\lambda r(\lambda)}, \lambda W) \in I\} = \mu_+(I).$$

Substituting W in the above by the scaled W_λ and then using Lemma 1.1 we obtain the following result.

<u>Theorem 3.1 ([9]).</u> (i) The full distribution of $(\log t)^{-2} X(t)$ converges to μ_+ as $t \to \infty$. (ii) μ_+ has a symmetric density and for $\lambda > 0$

$$\int_0^\infty e^{-\lambda x}\,\mu_+(dx) = \int_0^\infty \frac{\cosh\sqrt{2\lambda} - 1}{(\sigma+1)^3\{\cosh\sqrt{2\lambda} + \sigma\frac{\sinh\sqrt{2\lambda}}{\sqrt{2\lambda}}\}\lambda}\,d\sigma\ ,\ \lambda > 0.$$

4. Symmetric Stable Environment

In this section the space $\underline{\underline{W}}$ of the environments is $D(\mathbb{R} \to \mathbb{R})$[1] $\cap \{W(0) = 0\}$ and Q is the probability measure on $\underline{\underline{W}}$ with respect to which $\{W(u) : u>0\}$ and $\{W(-u) : u>0\}$ are independent and symmetric stable processes with exponent $\alpha(0 < \alpha < 2)$ such that

$$E^Q\{e^{\sqrt{-1}\xi W(u)}\} = e^{-|u||\tau|^\alpha}\ ,\ u\in \mathbb{R},\quad \xi\ \ \mathbb{R}.$$

This case is contained in the frame of Schumacher's work [7] (see also Kawazu-Tamura-Tanaka [4] for additional information) and so the full distribution of $(\log t)^{-\alpha} X(t,W)$ is convergent as $t \to \infty$. The purpose this section is to give a simple probabilistic description of the limit distribution. In the case of a Brownian environment Kesten [6] proved that the density of the limit distribution is given by

$$(4.1)\qquad \frac{2}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}\,e^{-(2k+1)^2\pi^2|x|/8}\ ,\ x\in \mathbb{R}.$$

We begin by stating some known results (see [7], [4]). First we give the definition of a valley, which is a slight modification of the one given in §1. Let $W\in \underline{\underline{W}}$. By definition W is said to be <u>oscillating</u> at x ($\in\mathbb{R}$) if

$$\sup_{I_\pm} W > W(x\pm)\ ,\ \inf_{I_\pm} W < W(x\pm)$$

for any $\varepsilon > 0$, where $I_+ = (x, x+\varepsilon)$ and $I_- = (x-\varepsilon, x)$. W is said to have a local minimum at x if $\inf_I W = W(x)\wedge W(x-)$ for some $\varepsilon>0$ where $I = (x-\varepsilon, x+\varepsilon)$. A local maximum is defined similarly. Denote by $\underline{\underline{W}}^\#$ the set of $W\in \underline{\underline{W}}$ with the following four properties.

1) This is the space of $\mathbb{R}$-valued right continuous functions on $\mathbb{R}$ with left limits.

(4.2a) $$\overline{\lim_{x\to\infty}}\{W(x) - \inf_{[0,x]} W\} = \overline{\lim_{x\to-\infty}}\{W(x) - \inf_{[x,0]} W\} = \infty .$$

(4.2b) If W is discontinuous at x, then W is oscillating at x.

(4.2c) For any open set G in $\mathbb{R}$ both the sets

$$\{x \in G : W(x) = \sup_G W\}, \ \{x \in G : W(x) = \inf_G W\}$$

contain at most one point.

(4.2d) W does not have a local maximum at $x = 0$.

Let $W \in \underline{\underline{W}}^{\#}$. Then $V = (a,b,c)$ is called a <u>valley</u> of W if

(i) $a < b < c$,

(ii) W is continuous at a, b and c,

(iii) $W(b) < W(x) < W(a)$ for every $x \in (a,b)$,

$W(b) < W(x) < W(c)$ for every $x \in (b,c)$,

(iv) $H_{a,b} < W(c) - W(b)$, $H_{c,b} < W(a) - W(b)$.

Here the notation $H_{x,y}$ is defined as in (1.2). The depth D and the inner directed ascent A are defined as in §1. Notice that (ii) implies that W is continuous at x if W has a local minimum (or maximum) at x. It can be proved that $Q(W^{\#}) = 1$. Moreover, making use of the self-similarity of a symmetric stable environment we can prove the following: There exists a valley $V = (a,b,c)$ of W such that $a < 0 < c$ and $A < 1 < D$, Q-a.s. The bottom of such a valley is uniquely determined by W and so is denoted by $b = b(W)$. For $\lambda > 0$ we define the scaled environment W_λ^α by $W_\lambda^\alpha(x) = \lambda^{-1}W(\lambda^\alpha x)$, $x \in \mathbb{R}$. Then as a special case of [6] (see also [4]) we have the following result: For any $\varepsilon>0$

$$P\{|\lambda^{-\alpha}X(e^\lambda, W) - b(W_\lambda^\alpha)| > \varepsilon\} \to 0$$

in probability with respect to Q as $\lambda\to\infty$ and consequently the full distribution of $(\log t)^{-\alpha}X(t)$ converges to the distribution of $b = b(W)$ as $t \to \infty$.

Our task is now to compute the distribution of $b = b(W)$. Although we are unable to give an analytic representation like (4.1), we have the following simple probabilistic representation of the limit distribution.

Theorem 4.1. The full distribution of $(\log t)^{-\alpha} X(t)$ converges, as $t \to \infty$, to the distribution with density

$$\phi(x) = \left\{ \frac{\Gamma(\alpha+1)}{\Gamma(\frac{\alpha}{2})} \right\}^2 \cdot Q\{ \sup_{[0,|x|]} W^* < 1 \} , \; x \in \mathbb{R},$$

where

$$W^*(u) = W(u) - \inf_{[0,u]} W, \; u > 0.$$

The proof will be based on the following description of $b(W)$ due to Kesten. For $W \in \underline{\underline{W}}^{\#}$ we set

$$d_{\pm} = \pm \inf\{u > 0 : W^*(\pm u) > 1\}$$

where $W^*(-u)$, $u>0$, is defined in a way similar to $W^*(u)$, and define $b_{\pm}$ by

$$W(b_+) = \inf_{[0,d_+]} W, \quad W(b_-) = \inf_{[d_-,0]} W.$$

We also set

$$V_{\pm} = W(b_{\pm}), \quad M_+ = \sup_{[0,b_+]} W, \quad M_- = \sup_{[b_-,0]} W.$$

Lemma 4.1 (Kesten [6]). Let (a, b, c) be a valley of W such that $a < 0 < c$ and $A < 1 < D$. Then

$$b = b_+ \text{ or } b_-$$

and the equality $b = b_+$ holds if and only if one of the following conditions holds:

(i) $V_- > V_+$ and $M_+ < (V_- + 1) \vee M_-$,

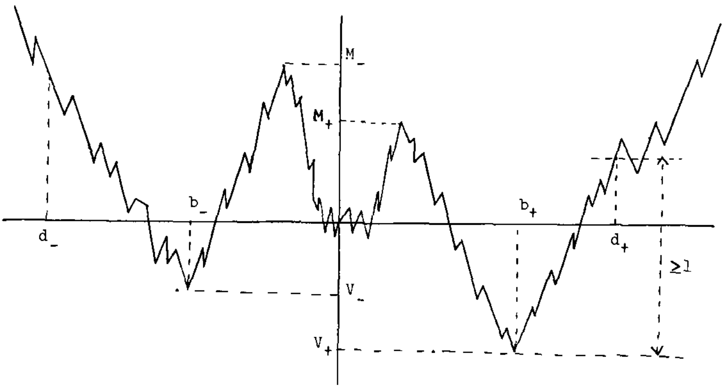
(ii) $V_- < V_+$ and $M_- > (V_+ + 1) \vee M_+$.

Since we deal only with the probability measure Q in the sequel, we write simply $E\{\cdot\}$ instead of $E^Q\{\cdot\}$ for the expectation with respect to Q. Again we introduce notation:

$$F_\lambda(x) = E\{e^{-\lambda b_+} \ ; \ V_+ < x - 1, \ M_+ < x\}$$

$$= E\{e^{-\lambda b_-} \ ; \ V_- < x - 1, \ M_- < x\} \ .$$

Clearly $F_\lambda(x)$ vanishes identically on $(-\infty, 0]$ and equals $E\{e^{-\lambda b_+}\}$ on $[1,\infty)$.



Lemma 4.2. $E\{e^{-\lambda b}; \ b > 0\} = \int_0^1 F_\lambda(x)dF_0(x) \ , \ \lambda > 0.$

Proof. Let

$$E_1 = \{V_- > V_+ \ , \ M_+ < (V_- + 1) \vee M_-\} \ ,$$

$$E_2 = \{V_1 < V_+ \ , \ M_- > (V_+ + 1) \vee M_+\} \ .$$

Then it is easy to see that

$$(E_1 \cup E_2)^c = \{V_- < V_+ \vee (M_+ - 1), \ M_- < (V_+ + 1) \vee M_+\},$$

and hence

$$E\{e^{-\lambda b} \ ; \ b > 0\}$$

$$= E\{e^{-\lambda b_+} \ ; \ V_- > V_+ \ , \ M_+ < (V_- + 1) \vee M_-\}$$

$$+ E\{e^{-\lambda b_+} \ ; \ V_- < V_+ \ , \ M_- > (V_+ + 1) \vee M_+\}$$

$$= E\{e^{-\lambda b_+} \ ; \ E_1 \cup E_2\}$$

$= E\{e^{-\lambda b_+}\} - E\{e^{-\lambda b_+}\ ;\ (E_1 \cup E_2)^c\}$

$= E\{e^{-\lambda b_+}\} - E\{e^{-\lambda b_+}\ ;\ V_- < V_+ \vee (M_+ - 1),\ M_- < (V_+ + 1) \vee M_+\}$

$= F_\lambda(1) - E\{e^{-\lambda b_+} \cdot F_0((V_+ + 1) \vee M_+)\}$

$= F_\lambda(1) - \int_0^1 F_0(x) dF_\lambda(x)$

$= \int_0^1 F_\lambda(x) dF_0(x).$

Now the rest of the proof of Theorem 4.1 is divided into two parts.

1. Properties of a symmetric stable process. We deal with the symmetric stable process $\{x + W(t),\ t \geq 0,\ x \in \mathbb{R},\ Q\}$ and prepare some of its properties for our later use. For the proof omitted here, see [10]. For $x,\ a \in \mathbb{R}$ we set

$$S_a^x = \inf\{t > 0:\ x + W(t) > a\},$$

$$T_a^x = \inf\{t > 0 :\ x + W(t) < a\},$$

$$T^x = S_1^x \wedge T_0^x .$$

Let $g_\lambda(x,y)$ be the Green function of order $\lambda > 0$ of the symmetric stable process with absorbing barriers at 0 and 1. We thus have

$$E\{\int_0^{T^x} e^{-\lambda t} f(x + W(t)) dt\} = \int_0^1 g_\lambda(x,y) f(y) dy,\quad 0 < x < 1.$$

It is known that $g_\lambda(x,y)$ is symmetric in x and y. We still need the following notation:

$$p_\lambda^+(x) = E\{e^{-\lambda S_1^x}\ ;\ S_1^x < T_0^x\},\quad p_\lambda^-(x) = E\{e^{-\lambda T_0^x}\ ;\ T_0^x < S_1^x\},$$

$$p_\lambda(x) = p_\lambda^+(x) + p_\lambda^-(x) = E\{e^{-\lambda T^x}\},$$

$$q^+(x) = 2\alpha^{-1}\{\Gamma(\alpha/2)\}^{-2}\, x^{\frac{\alpha}{2}} (1-x)^{\frac{\alpha}{2}-1},$$

$$q^-(x) = 2\alpha^{-1}\{\Gamma(\alpha/2)\}^{-2}\, (1-x)^{\frac{\alpha}{2}} x^{\frac{\alpha}{2}-1},$$

$$q^{\pm}_{\lambda}(x) = q^{\pm}(x) - \lambda \int_0^1 g_{\lambda}(x,y)q^{\pm}(y)dy,$$

$$r(x) = x^{\frac{\alpha}{2}-1}(1-x)^{\frac{\alpha}{2}-1}.$$

Notice that

$$g_{\lambda}(x,y) = g_{\lambda}(1-x,\ 1-y),$$

$$p^{+}_{\lambda}(x) = p^{-}_{\lambda}(1-x),\ p_{\lambda}(x) = p_{\lambda}(1-x),$$

$$q^{+}(x) = q^{-}(1-x)\ ,\ q^{+}_{\lambda}(x) = q^{-}_{\lambda}(1-x).$$

We also use the notation:

$$c(\alpha) = \Gamma(\alpha+1)\pi^{-1}\sin(\alpha\pi/2),$$

$$\langle p,\ q\rangle = \int_0^1 p(x)q(x)dx\ .$$

Lemma 4.3. $\alpha\{\Gamma(\alpha/2)\}^2 \langle p_{\lambda}\ ,\ q^{+}\rangle = \langle p_{\lambda}\ ,\ r\rangle$.

Proof. The left hand side equals

$$2\langle p_{\lambda},\ xr\rangle = 2\langle p_{\lambda}\ ,\ (1-x)r\rangle = \text{the right hand side};$$

the second equality is obtained by adding the preceding two terms and then dividing the sum by 2.

Lemma 4.4 (Watanabe [10]). (i) For any $f \in C([0,1])$

$$\lim_{x\downarrow 0} x^{-\alpha/2} \int_0^1 g_{\lambda}(x,y)f(y)dy = \langle f,q^{-}_{\lambda}\rangle,$$

$$\lim_{x\downarrow 0} x^{-\alpha/2} \int_0^1 g_{\lambda}(1-x,\ y)f(y)dy = \langle f,\ q^{+}_{\lambda}\rangle\ .$$

(ii) $$p^{+}_{\lambda}(x) = \alpha^{-1}c(\alpha) \int_0^1 g_{\lambda}(x,y)(1-y)^{-\alpha}\,dy,$$

$$p^{-}_{\lambda}(x) = \alpha^{-1}c(\alpha) \int_0^1 g_{\lambda}(x,y)y^{-\alpha}\,dy\ ,$$

$$p^{+}_{0}(x) = 2^{1-\alpha}\Gamma(\alpha)\{\Gamma(\alpha/2)\}^{-2} \int_{-1}^{-1+2x} (1-y^2)^{\frac{\alpha}{2}-1}\,dy\ .$$

Lemma 4.5. We have

$$\text{(i)}\quad \lim_{x\downarrow 0} p_\lambda^+(x)x^{-\alpha/2} = \lim_{x\downarrow 0} p_\lambda^-(1-x)x^{-\alpha/2} = K - \lambda\langle p_\lambda^-, q^+\rangle,$$

$$\text{(ii)}\quad \lim_{x\downarrow 0} \{1 - p_\lambda^-(x)\}x^{-\alpha/2} = \lim_{x\downarrow 0} \{1 - p_\lambda^+(1-x)\}x^{-\alpha/2} = K + \lambda\langle p_\lambda^+, q^+\rangle$$

where

$$K = \alpha^{-1}c(\alpha)\int_0^1 q^+(x)x^{-\alpha}dx = 2\Gamma(1-\tfrac{\alpha}{2})c(\alpha)/\Gamma(\alpha/2).$$

Proof. We give the proof of (ii). Since

$$\int_0^1 g_\lambda(x,y)dy = \lambda^{-1}\{1 - p_\lambda(x)\}$$

we have

$$\begin{aligned}\lim_{x\downarrow 0}\{1 - p_\lambda^-(x)\}x^{-\alpha/2} &= \lim_{x\downarrow 0}\{1 - p_\lambda(x)\}x^{-\alpha/2} + \lim_{x\downarrow 0} p_\lambda^+(x)x^{-\alpha/2} \\ &= \lambda\lim_{x\downarrow 0} x^{-\alpha/2}\int_0^1 g_\lambda(x,y)dy + \lim_{x\downarrow 0} p_\lambda^+(x)x^{-\alpha/2} \\ &= \langle\lambda + \alpha^{-1}c(\alpha)\cdot x^{-\alpha}, q_\lambda^+\rangle \\ &= K + \lambda\langle p_\lambda, q^+\rangle - \lambda\langle p_\lambda^-, q^+\rangle \\ &= K + \lambda\langle p_\lambda^+, q^+\rangle .\end{aligned}$$

2. Computation of $E\{e^{-\lambda b}; b > 0\}$.

Lemma 4.6. For $\lambda, \mu > 0$

$$E\{e^{-\lambda b_+ - \mu(d_+ - b_+)}\} = \{K - \mu\langle p_\mu^-, q^+\rangle\}\{K + \lambda\langle p_\lambda^+, q^+\rangle\}^{-1}.$$

In particular

$$E\{e^{-\lambda b_+}\} = K\{K + \lambda\langle p_\lambda^+, q^+\rangle\}^{-1},$$

$$E\{e^{-\lambda d_+}\} = \{K - \lambda\langle p_\lambda^-, q^+\rangle\}\{K + \lambda\langle p_\lambda^+, q^+\rangle\}^{-1}.$$

Proof. The proof is similar to that in [6]. For $\varepsilon>0$ we set

$$T^{(0)} = T_\varepsilon^{(0)} = 0,$$

$$T^{(n)} = T_\varepsilon^{(n)} = \inf \{t > T^{(n-1)} : W(t) - W(T^{(n-1)}) < -\varepsilon\}$$

$$S^{(n)} = S_\varepsilon^{(n)} = \inf\{t > T^{(n-1)} : W(t) - W(T^{(n-1)}) > 1-\varepsilon\},\ n = 1,2,\ldots$$

Then

$$E\{e^{-\lambda b_+ - \mu(d_+ - b_+)}\}$$

$$= \lim_{\varepsilon\downarrow 0} \sum_{n=1}^{\infty} E\{e^{-\lambda T^{(n)} - \mu(S^{(n+1)} - T^{(n)})};\ T^{(k)} < S^{(k)} \text{ for } 1 \le \forall k \le n \text{ and } S^{(n+1)} < T^{(n+1)}\}$$

$$= \lim_{\varepsilon\downarrow 0} \sum_{n=1}^{\infty} \{E\{e^{-\lambda T^0_{-\varepsilon}} ;\ T^0_{-\varepsilon} < S^0_{1-\varepsilon}\}\}^n\, E\{e^{-\mu S^0_{1-\varepsilon}};\ S^0_{1-\varepsilon} < T^0_{-\varepsilon}\}$$

$$= \lim_{\varepsilon\downarrow 0} \sum_{n=1}^{\infty} \{p_\lambda^-(\varepsilon)\}^n \cdot p_\lambda^+(\varepsilon).$$

If we set

$$K_1 = K + \lambda\langle p_\lambda^+ , q^+\rangle,\quad K_2 = K - \lambda\langle p_\lambda^- , q^+\rangle,$$

then

$$1 - p_\lambda^-(\varepsilon) \sim K_1\varepsilon^{\alpha/2},\quad p_\lambda^+(\varepsilon) \sim K_2\varepsilon^{\alpha/2} \quad \text{as} \quad \varepsilon\downarrow 0 .$$

Therefore we have

$$E\{e^{-\lambda b_+ - \mu(d_+ - b_-)}\} = \lim_{\varepsilon\downarrow 0} \sum_{n=1}^{\infty} (1 - K_1\varepsilon^{\alpha/2})^n K_2\varepsilon^{\alpha/2}$$

$$= K_2/K_1 = \{K - \lambda\langle p_\lambda^-, q^+\rangle\}\{K + \lambda\langle p_\lambda^+ , q^+\rangle\}^{-1},$$

completing the proof of the lemma.

For the proof of the theorem it is enough to prove the following lemma.

Lemma 4.7. For $\lambda > 0$

$$E\{e^{-\lambda b};\ b>0\} = \{\Gamma(\alpha + 1)/\Gamma(\alpha/2)\}^2 \int_0^\infty e^{-\lambda t} Q\{d_+ > t\}dt.$$

Proof. We first notice that

$$F_\lambda(x) = E\{e^{-\lambda b_+} ; V_+ < x-1, M_+ < x\}$$

$$= E\{e^{-\lambda T^0_{x-1}} ; T^0_{x-1} < T^0_x\}E\{e^{-\lambda b_+}\} = p^+_\lambda(x)E\{e^{-\lambda b_+}\} ,$$

$$F_0(x) = p^+_0(x) = 2^{1-\alpha}\Gamma(\alpha)\{\Gamma(\alpha/2)\}^{-2} \int_{-1}^{-1+2x} (1-y^2)^{\frac{\alpha}{2}-1} dy,$$

$$dF_0(x) = \Gamma(\alpha)\{\Gamma(\alpha/2)\}^{-2} r(x)dx .$$

We thus have

$$E\{e^{-\lambda b} ; b > 0\} = \int_0^1 F_\lambda(x)dF_0(x) \tag{4.3}$$

$$= E\{e^{-\lambda b_+}\} \int_0^1 p^+_\lambda(x)\, dF_0(x) = \Gamma(\alpha)\{\Gamma(\alpha/2)\}^{-2} E\{e^{-\lambda b_+}\}<p^+_\lambda , r>$$

$$= 2^{-1}\Gamma(\alpha)\{\Gamma(\alpha/2)\}^{-2}E\{e^{-\lambda b_+}\}<p_\lambda, r> .$$

On the other hand, using Lemma 4.3 we have

$$\int_0^\infty e^{-\lambda t}Q\{d_+ > t\}dt = \lambda^{-1}\{1 - E(e^{-\lambda d_+})\} = <p_\lambda, q^+>\{K + \lambda<p^+_\lambda, q^+>\}^{-1}$$

$$= \alpha^{-1}\{\Gamma(\alpha/2)\}^{-2} <p_\lambda, r>\{K + \lambda<p^+_\lambda, q^+>\}^{-1}$$

$$= (\alpha K)^{-1}\{\Gamma(\alpha/2)\}^{-2} E\{e^{-\lambda b_+}\}<p_\lambda, r> .$$

Combining this with (4.3) we obtain Lemma 4.7. This completes also the proof of Theorem 4.1.

Remark. The density ϕ in Theorem 4.1 is symmetric and

$$\int_0^\infty e^{-\lambda x}\phi(x)dx = \frac{\alpha\{\Gamma(\alpha)\}^2}{\{\Gamma(\alpha/2)\}^4} \cdot \frac{<p_\lambda , r>}{K + \lambda<p^+_\lambda, q^+>} , \lambda > 0.$$

References

[1] T. Brox, A one-dimensional diffusion process in a Wiener medium, to appear in Ann. Probab.

[2] A.O. Golosov, The limit distributions for random walks in random environments, Soviet Math. Dokl., 28 (1983), 19-22.

[3] K. Itô and H.P. McKean, Diffusion Processes and Their Sample Paths, Springer-Verlag, Berlin, 1965.

[4] K. Kawazu, Y. Tamura and H. Tanaka, One-dimensional diffusions and random walks in random media, in preparation.

[5] H. Kesten, M.V. Kozlov and F. Spitzer, A limit law for random walk in a random environment, Compositio Math., 30 (1975), 145-168.

[6] H. Kesten, The limit distribution of Sinai's random walk in random environment, to appear in Physica.

[7] S. Schumacher, Diffusions with random coefficients, Contemporary Mathematics (Particle Systems, Random Media and Large Deviations, ed. by R. Durrett), 41 (1985), 351-356.

[8] Y.G. Sinai, The limiting behavior of a one-dimensional random walk in a random medium, Theory of Probab. and its Appl., 27 (1982), 256-268.

[9] H. Tanaka, Limit distribution for 1-dimensional diffusion in a reflected Brownian medium, preprint, 1986.

[10] S. Watanabe, On stable processes with boundary conditions, J. Math. Soc. Japan, 14(1962), 170-198.

J. Fac. Sci. Univ. Tokyo
Sect. IA, Math.
34 (1987), 351-369.

Stochastic differential equation corresponding to the spatially homogeneous Boltzmann equation of Maxwellian and non-cutoff type

Dedicated to Professor Seizô Itô on his 60th birthday

By Hiroshi TANAKA

Introduction

It is known that a Markov process can be associated with a certain nonlinear equation of Boltzmann type ([2][3][4][5][7]). In the case of the spatially homogeneous Boltzmann equation of Maxwellian molecules, the associated Markov process was constructed by solving certain stochastic differential equation (abbreviated: SDE) based on a Poisson random measure ([7], see also [5][6]). The purposes of this paper are to simplify the proof of existence of solutions of the SDE of [7] by modifying the form of the SDE and also to give some remarks concerning the uniqueness of solutions.

We consider the Boltzmann equation of Maxwellian molecules:

$$(1) \qquad \frac{\partial u}{\partial t}=\int_{(0,\pi)\times(0,2\pi)\times\boldsymbol{R}^3}(u'u'_1-uu_1)Q(\theta)d\theta\hat{d}\epsilon dx_1, \quad t\geq 0, \quad x\in\boldsymbol{R}^3,$$

where $u=u(t,x)$, $u_1=u(t,x_1)$, $u'=u(t,x')$, $u'_1=u(t,x'_1)$ and $\hat{d}\epsilon=d\epsilon/2\pi$. $Q(\theta)$, $0<\theta<\pi$, is a positive function determined by the intermolecular repulsive force which is inversely proportional to the fifth power of their distance and has the property: $Q(\theta)\sim const.\,\theta^{-3/2}$, $\theta\downarrow 0$; so $\int_0^\pi Q(\theta)d\theta=\infty$ (non-cutoff) but

$$(\mathrm{I}) \qquad \int_0^\pi \theta Q(\theta)d\theta<\infty.$$

However, the special form of $Q(\theta)$ is not important in our methods and hence in this paper we assume that $Q(\theta)$ is an arbitrary nonnegative function satisfying only the condition (I) or even the following weaker one:

(II) $$\int_0^\pi \theta^2 Q(\theta)d\theta<\infty.$$

The two cases (I) and (II) are discussed separately.

A molecule with velocity x collides with a similar test molecule with velocity x_1; the post-collision velocities are denoted by x' and x_1', respectively. If $S(x, x_1)$ denotes the 2-dimensional sphere with center $(x+x_1)/2$ and radius $|x-x_1|/2$, then x' and x_1' are always on $S(x, x_1)$, or more precisely, $S(x', x_1')=S(x, x_1)$. Taking a spherical coordinate system on $S(x, x_1)$ with north pole x, denote by θ (resp. ϵ) the colatitude (resp. longitude) of x'. Then x' and x_1' can be regarded as functions of x, x_1, θ and ϵ. We set

$$a(x, x_1, \theta, \epsilon)=x'-x.$$

A probability measure valued function $u(t)$, $t>0$, is called a weak solution of (1) if

$$\frac{d}{dt}\langle u(t), \varphi\rangle=\langle u(t)\otimes u(t), K\varphi\rangle, \quad \varphi\in C_0^\infty(\boldsymbol{R}^3),$$

where $(K\varphi)(x, x_1)=\int_{(0,\pi)\times(0,2\pi)}\{\varphi(x')-\varphi(x)\}Q(\theta)d\theta\hat{d}\epsilon$ (see Appendix of [7]).

In [7] the following SDE was considered in connection with the Boltzmann equation (1) under the assumption (I):

$$(2)\quad X(t, \omega)=X(0, \omega) + \int_{(0,t]\times(0,\pi)\times(0,2\pi)\times(0,1)} a(X(s-, \omega), Y(s-, \alpha), \theta, \epsilon)N(dsd\theta d\epsilon d\alpha).$$

Here, $N(\cdot)$ is a Poisson random measure on $(0, \infty)\times(0, \pi)\times(0, 2\pi)\times(0, 1)$ with intensity measure $dsQ(\theta)\hat{d}\epsilon d\alpha$, and the solution process $\{X(t, \omega), t\geq 0\}$ is to be found on a basic probability space $\{\Omega, P\}$ under the condition that the process $\{Y(t, \alpha), t\geq 0\}$, defined on the probability space $\{(0, 1), d\alpha\}$ and describing the motion of a test molecule, is equivalent in law to $\{X(t, \omega), t\geq 0\}$. The relation between the Boltzmann equation (1) and the SDE (2) is that the probability distribution of $X(t)$ is a weak solution of (1) (general theory of SDE's including jump parts goes back to K. Itô [1]).

The modification we are making for the SDE (2) in proving existence theorem is as follows:

(i) $N(\cdot)$ is replaced by a Poisson random measure (again denoted by $N(\cdot)$) on $(0,\infty)\times(0,\pi)\times S^2\times\Omega_1$ with intensity measure $dsQ(\theta)d\theta\hat{d}\sigma dP_1$, where $\{\Omega_1, P_1\}$ is a copy of the basic probability space $\{\Omega, P\}$ and $\hat{d}\sigma$ is the uniform probability distribution on the 2-dimensional unit sphere S^2.

(ii) $Y(s-,\alpha)$ is replaced by $X(s-,\omega_1)$.

(iii) $a(x, x_1, \theta, \epsilon)$ is replaced by $b(x, x_1, \theta, \sigma)$ (the definition is given in § 1).

Thus in the case (I) the modified SDE can be written as

$$(3)\qquad X(t,\omega)=X(0,\omega)+\int_{S_t} b(X(s-,\omega), X(s-,\omega_1),\theta,\sigma)N(dsd\theta d\sigma d\omega_1),$$

where $S_t=(0,t]\times(0,\pi)\times S^2\times\Omega_1$ (in the case (II) the modified SDE is given by (3.3) in § 3). Advantage of the modified SDE (3) is that the new coefficient $b(x, x_1, \theta, \sigma)$ is Lipschitz continuous in (x, x_1) as an $L^1(\hat{d}\sigma)$-valued function for each fixed θ (see Lemma 1) and that the process describing the motion of a test molecule is exactly a copy of the solution process $X(t,\omega)$, and in fact, by virtue of these, (3) can be solved easily by using a routine iteration method. The proof of pathwise uniqueness for (3) is also easy.

Most of the discussions on the uniqueness in the law sense are essentially the same as the proof of Theorem 4.1 of [7] but they are somewhat simplified. In formulating the uniqueness in the law sense we further modify the SDE (3) as follows:

(i′) $\{\Omega_1, P_1\}$ is replaced by a probability space $\{\tilde{\Omega}, \tilde{P}\}$ which need not be a copy of $\{\Omega, P\}$.

(ii′) $X(s-,\omega_1)$ is replaced by $\tilde{X}(s,\tilde{\omega})$ which is an arbitrary measurable process defined on $\{\tilde{\Omega}, \tilde{P}\}$ such that it has the same distribution as the solution $X(s,\omega)$ for each s.

The uniqueness in the law sense is proved for this modified SDE so (in the case (I)) the solution process has the same law as the solution process of (3) (and also (2)).

Similar discussions in the case (II) are also given.

§ 1. L^1-Lipschitz continuity of $b(x, x_1, \theta, \sigma)$

Think of $S(x, x_1)$ as a celestial globe with north pole x and let $C(x, x_1, \theta)$ denote the circle on $S(x, x_1)$ with constant colatitude θ. Given

$\sigma\in S^2$, let $s(x, x_1, \sigma)=2^{-1}|x-x_1|\sigma+2^{-1}(x+x_1)$, let $M(x, x_1, \sigma)$ denote the meridian on $S(x, x_1)$ passing through $s(x, x_1, \sigma)$ and set

$$b_0(x, x_1, \theta, \sigma)=C(x, x_1, \theta)\cap M(x, x_1, \theta)=\text{a point on } S(x, x_1),$$
$$b(x, x_1, \theta, \sigma)=b_0(x, x_1, \theta, \sigma)-x.$$

Then for fixed $x, x_1\in \boldsymbol{R}^3$, $x\neq x_1$ and $\theta\in(0, \pi)$, $b_0(x, x_1, \theta, \sigma)$ is uniformly distributed on $C(x, x_1, \theta)$ as a random variable defined on the probability space $\{S^2, \hat{d}\sigma\}$. When $x=x_1$, we set $b(x, x_1, \theta, \sigma)=0$.

LEMMA 1. *For any* $x, x_1, y, y_1\in \boldsymbol{R}^3$ *and* $\theta\in(0, \pi)$,

$$\int_{S^2}|b(x, x_1, \theta, \sigma)-b(y, y_1, \theta, \sigma)|\hat{d}\sigma\leq const.\ \{|x-y|+|x_1-y_1|\}\theta, \tag{1.1}$$

where const. is independent of x, x_1, y, y_1 *and* θ.

PROOF. First we consider a special case.

(i) Special case: $S(x, x_1)=S(y, y_1)=S^2$.

In this case the integral on the left in (1.1) depends only on θ and the angle ξ $(0\leq\xi\leq\pi)$ between x and y. Therefore, it is enough to consider the case

$$x=(0, 0, 1),\ y=(0, \sin\xi, \cos\xi),\ x_1=-x,\ y_1=-y, \tag{1.2}$$

and prove that the integral on the left in (1.1) is dominated by const. $\theta\xi$. Let A be the rotation in $\boldsymbol{R}^3$ around the x^1-axis by the angle ξ. Then in the case (1.2) we have

$$b(y, y_1, \theta, \sigma)=A^{-1}b(x, x_1, \theta, A\sigma). \tag{1.3}$$

A point $\sigma\in S^2$ is expressed as $\sigma=(r, \sqrt{1-r^2}\cos\varphi, \sqrt{1-r^2}\sin\varphi)$ where $-1\leq r\leq 1$, $0\leq\varphi<2\pi$. We assume $0\leq r\leq 1$ for simplicity. We notice that $A\sigma=(r, \sqrt{1-r^2}\cos(\varphi+\xi), \sqrt{1-r^2}\sin(\varphi+\xi))$. Next, we define α and $\tilde{\alpha}$, respectively, by

$$\cos\alpha=\frac{r}{\sqrt{r^2+(1-r^2)\cos^2\varphi}}, \qquad \sin\alpha=\frac{\sqrt{1-r^2}\cos\varphi}{\sqrt{r^2+(1-r^2)\cos^2\varphi}}$$
$$\cos\tilde{\alpha}=\frac{r}{\sqrt{r^2+(1-r^2)\cos^2(\varphi+\xi)}}, \quad \sin\tilde{\alpha}=\frac{\sqrt{1-r^2}\cos(\varphi+\xi)}{\sqrt{r^2+(1-r^2)\cos^2(\varphi+\xi)}}.$$

Then

$$b(x, x_1, \theta, \sigma) = (\sin\theta\cos\alpha, \sin\theta\sin\alpha, \cos\theta - 1)$$
$$b(x, x_1, \theta, A\sigma) = (\sin\theta\cos\tilde{\alpha}, \sin\theta\sin\tilde{\alpha}, \cos\theta - 1),$$

and hence from (1.3)

$$b(y, y_1, \theta, \sigma) = (\sin\theta\cos\tilde{\alpha}, \sin\theta\sin\tilde{\alpha}\cos\xi + (\cos\theta - 1)\sin\xi, \\ -\sin\theta\sin\tilde{\alpha}\sin\xi + (\cos\theta - 1)\cos\xi).$$

Therefore

$$(1.4) \quad b(x, x_1, \theta, \sigma) - b(y, y_1, \theta, \sigma) \\ = (\sin\theta(\cos\alpha - \cos\tilde{\alpha}), \sin\theta(\sin\alpha - \sin\tilde{\alpha}\cos\xi) + (1 - \cos\theta)\sin\xi, \\ \sin\theta\sin\tilde{\alpha}\sin\xi - (1 - \cos\theta)(1 - \cos\xi)),$$

and hence

$$\int_{S^2} |b(x, x_1, \theta, \sigma) - b(y, y_1, \theta, \sigma)| \hat{d}\sigma \\ = \frac{1}{2\pi}\int_0^1\int_0^{2\pi} |b(x, x_1, \theta, \sigma) - b(y, y_1, \theta, \sigma)| dr d\varphi \leq \text{const.}\, \theta\xi$$

follows from (1.4) once we prove the following estimates.

$$(1.5) \quad \int_0^1\int_0^{2\pi} |\cos\alpha - \cos\tilde{\alpha}| dr d\varphi \leq \text{const.}\, \xi.$$

$$(1.6) \quad \int_0^1\int_0^{2\pi} |\sin\alpha - \sin\tilde{\alpha}\cos\xi| dr d\varphi \leq \text{const.}\, \xi.$$

The proof of (1.5) is as follows. Setting $f(r, \varphi, \xi) = \cos\alpha - \cos\tilde{\alpha}$ and $f_\xi(r, \varphi, \xi) = \partial f/\partial\xi$, we have

$$\int_0^1\int_0^{2\pi} |\cos\alpha - \cos\tilde{\alpha}| dr d\varphi = \int_0^1\int_0^{2\pi} \left|\int_0^\xi f_\xi(r, \varphi, \eta) d\eta\right| dr d\varphi \\ \leq \int_0^\xi d\eta \int_0^1\int_0^{2\pi} |f_\xi(r, \varphi, \eta)| dr d\varphi \leq \text{const.}\, \xi,$$

because

$$\int_0^1\int_0^{2\pi} |f_\xi(r, \varphi, \eta)| dr d\varphi = \int_0^1\int_0^{2\pi} r(1 - r^2)\{r^2 + (1 - r^2)\cos^2\varphi\}^{-3/2} |\cos\varphi\sin\varphi| dr d\varphi \\ = 4\int_0^1\int_0^1 r(1 - r^2)\{r^2 + (1 - r^2)x^2\}^{-3/2} x\, dr dx \\ \leq 4\int_0^1\int_0^1 rx(rx)^{-3/2} dr dx < \infty \quad (\text{use } r^2 + (1 - r^2)x^2 \geq rx).$$

As for (1.6), it is enough to prove

$$\int_0^1\int_0^{2\pi} |\sin\alpha - \sin\tilde{\alpha}|\, dr d\varphi \le \text{const.}\,\xi,$$

and for this it is also enough to prove that

$$(1.7) \qquad \int_0^1\int_0^{2\pi} |g_\xi(r, \varphi, \xi)|\, dr d\varphi = \text{indep. of } \xi < \infty,$$

where $g(r, \varphi, \xi) = \sin\alpha - \sin\tilde{\alpha}$ and $g_\xi(r, \varphi, \xi) = \partial g/\partial\xi$. But the left hand side of (1.7) is dominated by $I_1 + I_2$, where

$$\begin{aligned} I_1 &= \int_0^1\int_0^{2\pi} \sqrt{1-r^2}\{r^2 + (1-r^2)\cos^2\varphi\}^{-1/2}|\sin\varphi|\, dr d\varphi \\ &\le 4\int_0^1\int_0^1 \{r^2 + (1-r^2)x^2\}^{-1/2} dr dx < \infty \quad (\text{because } r^2 + (1-r^2)x^2 \ge rx). \\ I_2 &= \int_0^1\int_0^{2\pi} (1-r^2)^{3/2}\{r^2 + (1-r^2)\cos^2\varphi\}^{-3/2}\cos^2\varphi|\sin\varphi|\, dr d\varphi \\ &\le 4\int_0^1\int_0^1 \{r^2 + (1-r^2)x^2\}^{-3/2} x^2 dr dx < \infty. \end{aligned}$$

(ii) General case: Since

$$b(x, x_1, \theta, \sigma) = b\left(\frac{x-x_1}{2}, -\frac{x-x_1}{2}, \theta, \sigma\right) = \frac{|x-x_1|}{2} b(e_1, -e_1, \theta, \sigma),$$

$$e_1 = \frac{x-x_1}{|x-x_1|}, \quad e_2 = \frac{y-y_1}{|y-y_1|},$$

we have

$$\begin{aligned} &\int_{S^2} |b(x, x_1, \theta, \sigma) - b(y, y_1, \theta, \sigma)|\, \hat{d}\sigma \\ &\le \int_{S^2} \frac{|x-x_1|}{2} \cdot |b(e_1, -e_1, \theta, \sigma) - b(e_2, -e_2, \theta, \sigma)|\, \hat{d}\sigma \\ &\quad + \left|\frac{|x-x_1|}{2} - \frac{|y-y_1|}{2}\right| \cdot \int_{S^2} |b(e_2, -e_2, \theta, \sigma)|\, \hat{d}\sigma \\ &\le \text{const.}\, \frac{|x-x_1|}{2} \cdot |e_1 - e_2|\theta + \left|\frac{|x-x_1|}{2} - \frac{|y-y_1|}{2}\right|\theta \\ &\le \text{const.}\left\{\left|(x-x_1) - \frac{|x-x_1|}{|y-y_1|}(y-y_1)\right| + \Big||x-x_1| - |y-y_1|\Big|\right\}\theta, \end{aligned}$$

where we have used the result of case (i). Now (1.1) follows from the following trivial inequalities.

$$\Big| |x-x_1|-|y-y_1| \Big| \le |x-y|+|x_1-y_1| .$$

$$\Big| (x-x_1)-\frac{|x-x_1|}{|y-y_1|}(y-y_1) \Big| \le 2(|x-y|+|x_1-y_1|) .$$

§ 2. Stochastic differential equation—I

In this section we assume that $Q(\theta)$ satisfies the condition (I).

2.1. Existence theorem

We assume that a basic probability space $\{\Omega, \mathcal{F}, P\}$, equipped with a filtration $\{\mathcal{F}_t\}$ of increasing sub-σ-fields of $\mathcal{F}$, satisfies the following conditions.

(2.1) The σ-field $\mathcal{F}_0$ contains all P-negligible sets and is rich enough in the sense that, for any probability distribution μ in $\boldsymbol{R}^3$, there exists an $\mathcal{F}_0$-measurable $\boldsymbol{R}^3$-valued random variable with distribution μ.

(2.2) There exists an $\mathcal{F}_t$-adapted Poisson random measure $N(\cdot)$ on $(0, \infty)\times(0, \pi)\times S^2\times\Omega_1$ with intensity measure $dsQ(\theta)d\theta\hat{d}\sigma dP_1$ where $\{\Omega_1, P_1\}$ is a copy of $\{\Omega, P\}$.

A Poisson random measure $N(\cdot)$ on $(0, \infty)\times(0, \pi)\times S^2\times\Omega_1$ is said to be $\mathcal{F}_t$-adapted if, for each $t\ge 0$, $\mathcal{F}_t^0\subset\mathcal{F}_t$ and $\mathcal{F}_t$ is indepedent of $\mathcal{F}_\infty^t$, where $\mathcal{F}_t^0$ (resp. $\mathcal{F}_\infty^t$) is the smallest σ-field on Ω with respect to which the random variables $N(A)$, $A\in\mathcal{A}_t^0$ (resp. $A\in\mathcal{A}_\infty^t$), are measurable; here $\mathcal{A}_t^0$ (resp. $\mathcal{A}_\infty^t$) denotes the class of measurable subsets of $S_t=(0, t]\times(0, \pi)\times S^2\times\Omega_1$ (resp. $S_\infty^t=(t, \infty)\times(0, \pi)\times S^2\times\Omega_1$).

REMARK 1. The conditions (2.1) and (2.2) are not severe restrictions on $\{\Omega, \mathcal{F}, P\}$; in fact, it is easy to see that even the unit interval (0, 1) with the Lebesgue measure satisfies these conditions.

The SDE we are going to discuss is the following (=(3)):

$$(2.3) \qquad X(t, \omega)=X(0, \omega)+\int_{S_t} b(X(s-, \omega), X(s-, \omega_1), \theta, \sigma)N(dsd\theta d\sigma d\omega_1).$$

By a solution of (2.3) we mean an $\mathcal{F}_t$-adapted process $X(t, \omega)$, $t\ge 0$, which is right continuous and has left limits for almost all ω. $X(t, \omega)$, $t\ge 0$, is said to be integrable if $\int_0^T E|X(t,\omega)|dt<\infty,\ 0<{}^{\forall}T<\infty.$

THEOREM 1. *Let the condition* (I) *be satisfied and let* $X(0,\omega)$ *be a given* $\mathcal{F}_0$*-measurable random variable with* $E|X(0,\omega)|<\infty$. *Then there exists a unique integrable solution of* (2.3).

PROOF. We set $X_0(t,\omega)=X(0,\omega)$, $t\geq 0$, and define $X_n(t,\omega)$, $n\geq 1$, successively by

$$X_n(t,\omega)=X(0,\omega)+\int_{S_t} b(X_{n-1}(s-,\omega),X_{n-1}(s-,\omega_1),\theta,\sigma)N(dsd\theta d\sigma d\omega_1). \tag{2.4}$$

The stochastic integral is well-defined for each n by virtue of the estimate $|b(x,x_1,\theta,\sigma)|\leq|x-x_1|\theta/2$. By Lemma 1 we have

$$\begin{aligned} &E\Big\{\sup_{0\leq s\leq t}|X_{n+1}(s)-X_n(s)|\Big\}\\ &\leq E\Big\{\int_{S_t}|b(X_n(s-,\omega),X_n(s-,\omega_1),\theta,\sigma)\\ &\qquad\qquad -b(X_{n-1}(s-,\omega),X_{n-1}(s-,\omega_1),\theta,\sigma)|N(dsd\theta d\sigma d\omega_1)\Big\}\\ &\leq \text{const.}\,E\Big[\int_{S_t}\{|X_n(s-,\omega)-X_{n-1}(s-,\omega)|\\ &\qquad\qquad +|X_n(s-,\omega_1)-X_{n-1}(s-,\omega_1)|\}\theta dsQ(\theta)d\theta P_1(d\omega_1)\Big]\\ &\leq \text{const.}\int_0^t E|X_n(s,\omega)-X_{n-1}(s,\omega)|ds, \end{aligned}$$

and hence

$$\sum_{n=0}^{\infty}E\Big\{\sup_{0\leq s\leq t}|X_{n+1}(s)-X_n(s)|\Big\}\leq\sum_{n=0}^{\infty}\frac{c(c't)^n}{n!}<\infty$$

with some constants c and c'. Therefore

$$X(t,\omega)=\lim_{n\to\infty}X_n(t,\omega)$$

exists as a uniform convergence on each finite t-interval (a.s.); clearly $X(t,\omega)$ is an integrable solution of (2.3). To prove the uniqueness, let $X(t,\omega)$ and $Y(t,\omega)$ be any integrable solutions of (2.3). Then we have $E|X(t)-Y(t)|\leq\text{const.}\int_0^t E|X(s)-Y(s)|ds$ and hence $X(t)=Y(t)$, $t\geq 0$, a.s.

2.2. Uniqueness in the law sense

The uniqueness in Theorem 1 asserts that there is only one solution of (2.3) so far as the basic probability space, the initial value and the Poisson random measure are fixed. Different choices of the basic probability space etc. yield different solutions, but we can prove that their probability laws in the path space are the same provided that their initial distributions are the same. We prove this uniqueness in the law sense for a slightly modified SDE (Theorem 2).

Let $\{\Omega, \mathcal{F}, P\}$ be a probability space with a filtration $\{\mathcal{F}_t\}$ satisfying (2.1) and (2.2) as before. But now we replace $\{\Omega_1, P_1\}$ by $\{\tilde{\Omega}, \tilde{P}\}$ which need not be a copy of $\{\Omega, P\}$. A process $Y(t, \tilde{\omega})$ defined on $\{\tilde{\Omega}, \tilde{P}\}$ is said to be integrable if it is jointly measurable and if

$$\int_0^T \int_{\tilde{\Omega}} |Y(t, \tilde{\omega})| dt d\tilde{P} < \infty, \quad 0 < {}^\forall T < \infty.$$

Let W denote the space of $\boldsymbol{R}^3$-valued right continuous paths with left limits.

Given an integrable process $Y(t, \tilde{\omega})$ defined on $\{\tilde{\Omega}, \tilde{P}\}$, we consider the SDE

$$X(t, \omega) = X(0, \omega) + \int_{S_t} b(X(s-, \omega), Y(s, \tilde{\omega}), \theta, \sigma) N(ds d\theta d\sigma d\tilde{\omega}) \tag{2.5}$$

where $S_t = (0, t] \times (0, \pi) \times S^2 \times \tilde{\Omega}$.

PROPOSITION 1. *Let the condition* (I) *be satisfied and let* $X(0, \omega)$ *be a given* $\mathcal{F}_0$*-measurable random variable with* $E|X(0, \omega)| < \infty$. *Then for any given integrable process* $Y(t, \tilde{\omega})$ *there exists an integrable solution of* (2.5). *Also the law uniqueness holds in the following sense: The probability measure on* W *induced by a solution of* (2.5) *is uniquely determined by* u_0 *and* $\tilde{u}(t)$, $t \geq 0$, *where* u_0 *is the probability distribution of* $X(0, \omega)$ *and* $\tilde{u}(t)$ *is that of* $Y(\cdot, \tilde{\omega})$ *at time* t.

PROOF. The existence of a solution is proved by a routine iteration method as in the proof of Theorem 1. The law uniqueness is proved as follows. First we choose a sequence $\{h_n(t), t \geq 0\}$ of step functions such that

(2.6) each $h_n(t)$ is expressed as

$$h_n(t) = \begin{cases} 0 & \text{for} \quad t = 0 \\ t_{nk} & \text{for} \quad t_{nk} < t \leq t_{nk+1} \ (k = 0, 1, \cdots), \end{cases}$$

where $\{t_{nk}\}$ satisfies
$0=t_{n0}<t_{n1}<\cdots,\ \lim_{k\to\infty} t_{nk}=\infty,\ \lim_{n\to\infty}\sup_k(t_{nk+1}-t_{nk})=0;$

$$\lim_{n\to\infty}\int_0^t\int_{\tilde{\Omega}}|Y(s,\tilde{\omega})-Y(h_n(s),\tilde{\omega})|dsd\tilde{P}=0,\quad 0<{}^\forall t<\infty. \tag{2.7}$$

Let $X(t)$ be the solution of (2.5) and let $X_n(t)$ be the solution of

$$X_n(t)=X(0)+\int_{S_t} b(X_n(h_n(s)),\ Y(h_n(s),\tilde{\omega}),\theta,\sigma)N(dsd\theta d\sigma d\tilde{\omega}). \tag{2.8}$$

Then $X_n(t)$ is obtained as follows:

$$X_n(t)=X_n(t_{nk})+\int_{S_t-S_{t_{nk}}} b(X_n(t_{nk}),\ Y(t_{nk},\tilde{\omega}),\theta,\sigma)dN, \tag{2.9}$$
$$t_{nk}<t\leq t_{nk+1}\quad (k\geq 0).$$

Making use of the estimate

$$|b(x,x_1,\theta,\sigma)|\leq|x-x_1|\theta/2, \tag{2.10}$$

and also (2.7), we can easily prove that

$$E|X_n(s)|\leq \text{const.},\quad 0\leq s\leq t, \tag{2.11}$$

$$\sup\{E|X_n(t_1)-X_n(t_2)|:0\leq t_1,t_2\leq t,|t_1-t_2|\leq\varepsilon,n\geq 1\}\to 0\quad \text{as } \varepsilon\downarrow 0, \tag{2.12}$$

where const. may depend on t but not on n. Then, making use of Lemma 1 and then (2.7), (2.12), we have

$$\begin{aligned}E|X_n(t)-X(t)|&\leq \text{const.}\int_0^t E|X_n(h_n(s))-X(s)|ds\\&\quad+\text{const.}\int_0^t\int_{\tilde{\Omega}}|Y(h_n(s))-Y(s)|dsd\tilde{P}\\&\leq \text{const.}\int_0^t E|X_n(s)-X(s)|ds+o(1),\end{aligned}$$

and hence by Gronwall's inequality

$$\lim_{n\to\infty} E|X_n(t)-X(t)|=0. \tag{2.13}$$

On the other hand, by (2.9) we have for any $\xi\in\boldsymbol{R}^3$ and $t_{nk}<t\leq t_{nk+1}$

$$E\left[\exp\left\{\sqrt{-1}\xi\cdot X_n(t)\right\}\Big|\mathcal{F}_{t_{nk}}\right]$$

$$= \exp\left[\sqrt{-1}\xi\cdot x+(t-t_{nk})\int_{(0,\pi)\times S^2\times\tilde{\Omega}}\left\{e^{\sqrt{-1}\xi\cdot b(x,Y(t_{nk},\tilde{\omega}),\theta,\sigma)}-1\right\}Q(\theta)d\theta\hat{d}\sigma d\tilde{P}\right]$$
$$= \exp\left[\sqrt{-1}\xi\cdot x+(t-t_{nk})\int_{(0,\pi)\times S^2\times R^3}\left\{e^{\sqrt{-1}\xi\cdot b(x,y,\theta,\sigma)}-1\right\}Q(\theta)d\theta\hat{d}\sigma\tilde{u}(t_{nk},dy)\right]$$

where we put $x=X_n(t_{nk})$. This conditional expectation formula implies that the probability measure on W induced by the process $X_n(t)$ is uniquely determined by u_0 and $\tilde{u}(t)$, $t\geq 0$. Therefore, by (2.13) the probability measure on W induced by $X(t)$ is also uniquely determined by u_0 and $\tilde{u}(t)$, $t\geq 0$. This completes the proof of the proposition.

Now we consider the following SDE for which we are going to prove the law uniqueness:

$$X(t)=X(0)+\int_{S_t} b(X(s-),\tilde{X}(s,\tilde{\omega}),\theta,\sigma)N(dsd\theta d\sigma d\tilde{\omega}). \tag{2.14a}$$

Here, an $\mathcal{F}_t$-adapted integrable solution $X(t)$ is found under the condition that

(2.14b) $\tilde{X}(t,\tilde{\omega})$ is a measurable process defined on the probability space $\{\tilde{\Omega},\tilde{P}\}$ such that $\tilde{X}(t,\tilde{\omega})$ has the same distribution as $X(t)$ for each t.

REMARK 2. When $\{\tilde{\Omega},\tilde{P}\}=\{\Omega_1,P_1\}$, a solution of (2.3) is also a solution of (2.14).

THEOREM 2. *Let the condition* (I) *be satisfied and let* $X(0,\omega)$ *be any* R^3*-valued and* $\mathcal{F}_0$*-measurable random variable with* $E|X(0,\omega)|<\infty$. *Then the probability measure on* W *induced by any integrable solution of* (2.14) *is uniquely determined by the probability distribution* u_0 *of the initial value* $X(0,\omega)$.

PROOF. Let $\hat{\Omega}=[0,1]$, $\hat{\mathcal{F}}=$the σ-field of Borel subsets of $[0,1]$, $\hat{P}(A)=$ the Lebesgue measure of A $(\in\hat{\mathcal{F}})$ and let $\{\Omega_1,P_1\}$ be a copy of $\{\hat{\Omega},\hat{P}\}$. As in 2.1 we construct, on the probability space $\{\hat{\Omega},\hat{P}\}$, a u_0-distributed random variable $\hat{X}$ and a Poisson random measure $\hat{N}(\cdot)$ on $(0,\infty)\times(0,\pi)\times S^2\times\Omega_1$ with intensity measure $dtQ(\theta)d\theta\hat{d}\sigma dP_1$ so that $\hat{X}$ and $\hat{N}(\cdot)$ are independent. We then consider the SDE of the type (2.3)

$$\hat{X}(t)=\hat{X}+\int_{\hat{S}_t} b(\hat{X}(s-),\hat{X}(s-,\omega_1),\theta,\sigma)d\hat{N}, \tag{2.15}$$

where $\hat{S}_t=(0,t]\times(0,\pi)\times S^2\times\Omega_1$. We are going to prove that for any solution $X(t)$ of (2.14) there exists a solution of (2.15) which (as a process) is equivalent in law to $X(t)$. Once this has been proved, the law uniqueness of solutions of (2.14) follows immediately from the pathwise uniqueness of solutions of (2.15).

On the probability space $\{\Omega_1, P_1\}$ we can find an $\boldsymbol{R}^3$-valued right continuous process $\hat{X}_0(t,\omega_1)$ having left limits which is equivalent in law to a solution $X(t)$ of (2.14). Given such a process $\hat{X}_0(t,\omega_1)$, we consider the SDE

$$\hat{X}(t)=\hat{X}+\int_{\hat{S}_t} b(\hat{X}(s-),\hat{X}_0(s-,\omega_1),\theta,\sigma)d\hat{N}. \tag{2.16}$$

Since the both (test) processes $\tilde{X}(t,\tilde{\omega})$ and $\hat{X}_0(t-,\omega_1)$ in (2.14) and (2.16) have the same marginal distribution at each time t, Proposition 1 implies that the unique solution $\hat{X}_1(t)$ of (2.16) is equivalent in law to a solution process $X(t)$ of (2.14). Next we construct $\hat{X}_n(t)$ for $n\geq 2$ by $\hat{X}_n(t)=$the solution of (2.16) with $\hat{X}_0(s-,\omega_1)$ replaced by $\hat{X}_{n-1}(s-,\omega_1)$. Then as in the proof of Theorem 1 we can prove that $\hat{X}_n(t)$ converges to a solution $\hat{X}(t)$ of (2.15) as $n\to\infty$. Since each process $\hat{X}_n(t)$ is equivalent in law to $X(t)$, so is $\hat{X}(t)$. This completes the proof of Theorem 2.

§ 3. Stochastic differential equation—II

In this section we assume

$$\int_0^\pi \theta Q(\theta)d\theta=\infty, \qquad \int_0^\pi \theta^2 Q(\theta)d\theta<\infty. \tag{3.1}$$

Let $\{\Omega, P\}$, $\{\Omega_1, P_1\}$ and $N(\cdot)$ be the same as in 2.1 and set $M(A)=N(A)-\lambda(A)$ for a measurable subset A of $(0,\infty)\times(0,\pi)\times S^2\times\Omega_1$ with $\lambda(A)=\int_A dtQ(\theta)d\theta\hat{d}\sigma dP_1<\infty$. Then the stochastic integral on the right of (2.3) can be written as

$$\int_{S_t} b(X(s-,\omega),X(s-,\omega_1),\theta,\sigma)dM+\int_{S_t} b(X(s-,\omega),X(s-,\omega_1),\theta,\sigma)d\lambda.$$

The first integral in the above will make sense under the condition (3.1) while the second integral equals

$$-c\int_0^t \{X(s,\omega)-\overline{X(s,\omega)}\}ds$$

where $\overline{X(s, \omega)}=E\{X(s, \omega)\}$ and $c=\int_0^\pi 2^{-1}(1-\cos\theta)Q(\theta)d\theta$. So we are led to the following SDE:

$$(3.2)\qquad X(t, \omega)=X(0, \omega)+\int_{S_t} b(X(s-, \omega), X(s-, \omega_1), \theta, \sigma)dM \\ -c\int_0^t \{X(s, \omega)-\overline{X(s, \omega)}\}ds.$$

If $\int_{S^2} |b(x, x_1, \theta, \sigma)-b(y, y_1, \theta, \sigma)|^2\hat{d}\sigma$ were dominated by a constant multiple of $\{|x-y|^2+|x_1-y_1|^2\}\theta^2$, we could solve (3.2) easily. But this is not likely to be true. So we further modify the SDE (3.2) so that it can be solved in an easier way. First we introduce the predictable σ-field $\mathscr{P}$ on $[0, \infty)\times\Omega\times\Omega_1$; it is defined as the smallest σ-field on $[0, \infty)\times\Omega\times\Omega_1$ with respect to which all functions $a(t, \omega, \omega_1)$ satisfying the following conditions (i) and (ii) are measurable.

(i) For each fixed $t\geq 0$, $a(t, \omega, \omega_1)$ is $\mathscr{F}_t\otimes\hat{\mathscr{F}}$-measurable where $\{\Omega_1, \hat{\mathscr{F}}, P_1\}$ is a copy of $\{\Omega, \mathscr{F}, P\}$.

(ii) For fixed ω and ω_1, $a(t, \omega, \omega_1)$ is left continuous in t.

Let $\mathscr{R}$ denote the class of predictable processes (i.e., $\mathscr{P}$-measurable functions on $[0, \infty)\times\Omega\times\Omega_1$) with values in the space 0(3) of orthogonal matrices of degree 3. Then our modified SDE can be written as

$$(3.3)\quad X(t, \omega)=X(0, \omega)+\int_{S_t} b(X(s-, \omega), X(s-, \omega_1), \theta, R(s, \omega, \omega_1)\sigma)dM \\ -c\int_0^t \{X(s, \omega)-\overline{X(s, \omega)}\}ds.$$

By a solution of (3.3) we mean an $\mathscr{F}_t$-adapted process $X(t, \omega)$, $t\geq 0$, which is right continuous in t, has left limits for almost all ω and satisfies (3.3) with some $R=R(t, \omega, \omega_1)\in\mathscr{R}$. $X(t, \omega)$ is said to be square integrable if $\int_0^T E\{|X(t, \omega)|^2\}dt<\infty$, $0<{}^\forall T<\infty$.

REMARK 3. If we set

$$\tilde{N}(A)=\int_{S_\infty} \mathbf{1}_A(s, \theta, R(s, \omega, \omega_1)\sigma)N(dsd\theta d\sigma d\omega_1), \\ \tilde{M}(A)=\tilde{N}(A)-\lambda(A),$$

then $\tilde{N}(\cdot)$ is also an $\mathscr{F}_t$-adapted Poisson random measure with the same intensity measure λ and (3.3) becomes (3.2) with M replaced by $\tilde{M}$. In this sense (3.2) and (3.3) may be regarded as equivalent. The corresponding martingale problems are the same.

For $\sigma, \sigma' \in S^2$ we denote by $R(\sigma, \sigma')$ the rotation (orthogonal matrix) in $\mathbf{R}^3$ which sends σ to σ' along the geodesic connecting σ with σ', and set $R(x, x_1, y, y_1) = R((x-x_1)|x-x_1|^{-1}, (y-y_1)|y-y_1|^{-1})$ (for $x \neq x_1, y \neq y_1$), $=$the identity matrix (otherwise). Then we have the following lemma (see Lemma 3.1 of [7]).

LEMMA 2. *For any* $x, x_1, y, y_1 \in \mathbf{R}^3$

$$|b(x, x_1, \theta, \sigma) - b(y, y_1, \theta, R(x, x_1, y, y_1)\sigma)| \leq \text{const.}\{|x-y| + |x_1-y_1|\}\theta.$$

THEOREM 3. *Let the condition* (II) *be satisfied and let* $X(0, \omega)$ *be a given* $\mathcal{F}_0$*-measurable random variable with* $E\{|X(0, \omega)|^2\} < \infty$. *Then there exists a square integrable solution of* (3.3). *Moreover, the law uniqueness holds for* (3.3) *in the sense that the probability measure on* W *induced by any square integrable solution of* (3.3) *is uniquely determined by the probability distribution* u_0 *of* $X(0, \omega)$.

PROOF. Define $X_n(t, \omega)$, $n \geq 0$, by

$$\begin{aligned} X_0(t, \omega) &= X(0, \omega), \\ X_n(t, \omega) &= X(0, \omega) + \int_{S_t} b(X_{n-1}(s-, \omega), X_{n-1}(s-, \omega_1), \theta, R_{n-1}\sigma) dM \\ &\quad - c\int_0^t \{X_{n-1}(s, \omega) - \overline{X_{n-1}(s, \omega)}\} ds, \qquad n \geq 1, \end{aligned}$$

where $R_n = R_n(s, \omega, \omega_1) = \prod_{k=1}^{n} R(X_{k-1}(s-, \omega), X_{k-1}(s-, \omega_1), X_k(s-, \omega), X_k(s-, \omega_1))$. Then, making use of Lemma 2 and the convergence of the second integral of (3.1) we have

$$\begin{aligned} & E\Big\{\sup_{0 \leq s \leq t} |X_{n+1}(s) - X_n(s)|^2\Big\} \\ & \leq 8E\Big\{\int_{S_t} |b(X_n(s-, \omega), X_n(s-, \omega_1), \theta, R_n\sigma) - b(X_{n-1}(s-, \omega), X_{n-1}(s-, \omega_1), \\ & \qquad \theta, R_{n-1}\sigma)|^2 d\lambda\Big\} \\ & \quad + 2c^2 t\int_0^t E\{|X_n(s, \omega) - \overline{X_n(s, \omega)} - X_{n-1}(s, \omega) + \overline{X_{n-1}(s, \omega)}|^2\} ds \\ & \leq \text{const.}\ (1+t)\int_0^t E\{|X_n(s) - X_{n-1}(s)|^2\} ds, \end{aligned}$$

and hence by a routine argument we can prove that

$$\sum_{n=1}^{\infty} \sup_{0\le s\le t} |X_n(s, \omega) - X_{n-1}(s, \omega)| \tag{3.4}$$

is convergent for all $t\geq 0$ with probability 1. We denote by $\tilde{\Omega}$ the set of ω for which (3.4) is convergent for all $t\geq 0$ and also by $\tilde{\Omega}_1$ $(\subset\Omega_1)$ the copy of $\tilde{\Omega}$. We set

$$X(t, \omega) = \begin{cases} \lim_{n\to\infty} X(t, \omega), & \omega\in\tilde{\Omega} \\ 0 \quad , & \text{otherwise,} \end{cases}$$

$$\Gamma = \{(t, \omega, \omega_1) \in [0, \infty)\times\tilde{\Omega}\times\tilde{\Omega}_1 \; : \; X(t-, \omega) \neq X(t-, \omega_1)\}.$$

Then $\Gamma\in\mathscr{P}$. We first claim that $R_n(t, \omega, \omega_1)$ is convergent as $n\to\infty$ for each fixed $(t, \omega, \omega_1)\in\Gamma$. Using the notation $\|A\|=\sup\{|Ax| : |x|=1\}$ for a matrix A, we have

$$\begin{aligned}
&\|R_n(t, \omega, \omega_1) - R_{n-1}(t, \omega, \omega_1)\| \\
&= \|\{R(X_{n-1}(t-, \omega), X_{n-1}(t-, \omega_1), X_n(t-, \omega), X_n(t-, \omega_1)) - I\}R_{n-1}(t, \omega, \omega_1)\| \\
&\le \sqrt{3}\, \|R(X_{n-1}(t-, \omega), X_{n-1}(t-, \omega_1), X_n(t-, \omega), X_n(t-, \omega_1)) - I\| \\
&\le \sqrt{3} \left| \frac{X_{n-1}(t-, \omega) - X_{n-1}(t-, \omega_1)}{|X_{n-1}(t-, \omega) - X_{n-1}(t-, \omega_1)|} - \frac{X_n(t-, \omega) - X_n(t-, \omega_1)}{|X_n(t-, \omega) - X_n(t-, \omega_1)|} \right| \\
&\le \frac{2\sqrt{3}\{|X_{n-1}(t-, \omega) - X_n(t-, \omega) + X_{n-1}(t-, \omega_1) - X_n(t-, \omega_1)|\}}{|X_n(t-, \omega) - X_n(t-, \omega_1)|}.
\end{aligned}$$

If $(t, \omega, \omega_1)\in\Gamma$, then $|X_n(t-, \omega) - X_n(t-, \omega_1)|\geq\varepsilon$ for some $\varepsilon>0$, and for all sufficiently large n (say, for $n\geq n_0$) and hence

$$\begin{aligned}
&\sum_{n=n_0}^{\infty} \|R_n(t, \omega, \omega_1) - R_{n-1}(t, \omega, \omega_1)\| \\
&\le \frac{2\sqrt{3}}{\varepsilon} \sum_{n=n_0}^{\infty} \{|X_n(t-, \omega) - X_{n-1}(t-, \omega)| + |X_n(t-, \omega_1) - X_{n-1}(t-, \omega_1)|\} < \infty.
\end{aligned}$$

Next we define $R\in\mathscr{R}$ by

$$R(t, \omega, \omega_1) = \begin{cases} \lim_{n\to\infty} R_n(t, \omega, \omega_1), & (t, \omega, \omega_1)\in\Gamma \\ \text{identity}, & \text{otherwise,} \end{cases}$$

and claim that $\{X(t, \omega), R\}$ is a solution of (3.3). For this it is enough to prove that

$$E\left[\left|\int_{S_t} \{b(X_n(s-, \omega), X_n(s-, \omega_1), \theta, R_n\sigma) - b(X(s-, \omega), X(s-, \omega_1), \theta, R\sigma)\} dM\right|^2\right] \tag{3.5}$$

tends to 0 as $n\to\infty$. If $\tilde{R}_n=R(X_n(s-,\omega), X_n(s-,\omega_1), X(s-,\omega), X(s-,\omega_1))$, then (3.5) is dominated by

$$2E\left\{\int_{S_t}|b(X_n(s-,\omega), X_n(s-,\omega_1), \theta, R_n\sigma)\right.$$
$$\left.-b(X(s-,\omega), X(s-,\omega_1), \theta, \tilde{R}_nR_n\sigma)|^2 d\lambda\right\}$$
$$+2E\left\{\int_{S_t}|b(X(s-,\omega)-X(s-,\omega_1), \theta, \tilde{R}_nR_n\sigma)\right.$$
$$\left.-b(X(s-,\omega), X(s-,\omega_1), \theta, R\sigma)|^2 d\lambda\right\}$$
$$\leq \text{const.}\int_0^t E\{|X_n(s)-X(s)|^2\}ds+2E\left\{\int_{S_t}b_n(s,\omega,\omega_1,\theta,\sigma)d\lambda\right\},$$

where

$$b_n(s,\omega,\omega_1,\theta,\sigma)=|b(X(s-,\omega), X(s-,\omega_1), \theta, R_n'\sigma)$$
$$-b(X(s-,\omega), X(s-,\omega_1), \theta, \sigma)|^2, \quad R_n'=\tilde{R}_nR_nR^{-1}.$$

Moreover, we can easily prove that $R_n'(s,\omega,\omega_1)\to I$ as $n\to\infty$ for each fixed $(s,\omega,\omega_1)\in\Gamma$, and hence for each fixed $(s,\omega,\omega_1,\theta)\in(0,\infty)\times\tilde{\Omega}\times\tilde{\Omega}_1\times(0,\pi)$ we have

$$\lim_{n\to\infty}b_n(s,\omega,\omega_1,\theta,\sigma)=0$$

for almost all σ with respect to $\hat{d}\sigma$. Since we also have the bound $b_n(s,\omega,\omega_1,\theta,\sigma)\leq|X(s-,\omega)-X(s-,\omega_1)|^2\cdot\theta^2$, an application of Lebesgue's dominated convergence theorem yields

$$\lim_{n\to\infty}E\left\{\int_{S_t}b_n(s,\omega,\omega_1,\theta,\sigma)d\lambda\right\}=0,$$

which implies that (3.5) tends to 0 as $n\to\infty$. Thus the existence proof is finished.

To prove the law uniqueness let $h_n(t)=k2^{-n}$ for $k2^{-n}<t\leq(k+1)2^{-n}$ and $h(0)=0$. Let $X(t,\omega)$ be any square integrable solution of (3.3) with auxiliary process $R=R(t,\omega,\omega_1)\in\mathcal{R}$ and consider the SDE

$$Y_n(t,\omega)=X(0,\omega)+\int_{S_t}b(Y_n(h_n(s),\omega), Y_n(h_n(s),\omega_1), \theta, R_n\sigma)dM \tag{3.6}$$
$$-c\int_0^t\{Y_n(h_n(s))-\overline{Y_n(h_n(s))}\}ds,$$

where

$$R_n=R_n(s,\omega,\omega_1)$$
$$=R(X(s-,\omega), X(s-,\omega_1), Y_n(h_n(s),\omega), Y_n(h_n(s),\omega_1))R(s,\omega,\omega_1).$$

(3.6) can be solved easily; in fact, once we know $Y_n(t,\omega)$ for $0\le t\le k2^{-n}$, we can define $Y_n(t,\omega)$ for $k2^{-n}<t\le(k+1)2^{-n}$ by the right hand side of (3.6). We can prove that

$$E\{|Y_n(t)-Y_n(s)|^2\}\le \text{const.}\,|t-s|,\quad 0\le s<t\le T,$$

and also, by making use of Lemma 2, that

$$E\{|Y_n(t)-X(t)|^2\}\le \text{const.}\int_0^t E\{|Y_n(s)-X(s)|^2\}ds+\text{const.}\,\Delta_n(t),$$

where

$$\Delta_n(t)=\sup\{E(|Y_n(u)-Y_n(s)|^2):0\le s<u\le t,\,u-s\le 2^{-n}\}\le \text{const.}\,2^{-n},$$

const. being independent of n. Therefore

$$(3.7)\qquad E\{|Y_n(t)-X(t)|^2\}\longrightarrow 0,\qquad n\to\infty.$$

On the other hand, let $k2^{-n}<t\le(k+1)2^{-n}$, $x\in \boldsymbol{R}^3$ and set

$$\begin{aligned}\Phi_n(t)=ix\cdot Y_n(k2^{-n},\omega)+\int_{S_t-S_{k2^{-n}}}&(e^{ix\cdot b}-1-ix\cdot b)d\lambda\\ &-ic(t-k2^{-n})x\cdot\{Y_n(k2^{-n},\omega)-\overline{Y_n(k2^{-n},\omega)}\}\\ &\quad(\text{where } b=b(Y_n(k2^{-n},\omega),\,Y_n(k2^{-n},\omega_1),\theta,\sigma)).\end{aligned}$$

Then, for $k2^{-n}<t\le(k+1)2^{-n}$ we have

$$E\{e^{ix\cdot Y_n(t)}|\mathcal{F}_{k2^{-n}}\}=e^{\Phi_n(t)},\qquad \text{a.s.},$$

and hence the probability measure on W induced by the process $Y_n(t,\omega)$, $t\ge 0$, is uniquely determined by u_0. This combined with (3.7) proves the law uniqueness of square integrable solutions of (3.3).

In the rest of this section let $\{\Omega,P\}$ and $\{\tilde\Omega,\tilde P\}$ be the same as in 2.2 and consider the SDE

$$\begin{aligned}(3.8)\qquad X(t)=X(0)+\int_{S_t}&b(X(s-),\,Y(s,\tilde\omega),\theta,R(s,\omega,\tilde\omega)\sigma)dM\\ &-c\int_0^t\{X(s)-\overline{Y(s,\tilde\omega)}\}ds\end{aligned}$$

where $Y(t,\tilde\omega)$ is a given square integrable process defined on $\{\tilde\Omega,\tilde P\}$ and $R=R(t,\omega,\tilde\omega)$ is similar to one in (3.3).

PROPOSITION 2. *Let the condition* (II) *be satisfied and* $X(0,\omega)$ *be a given* $\mathcal{F}_0$*-measurable random variable with* $E\{|X(0,\omega)|^2\}<\infty$. *Then for* *any given square integrable process* $Y(t, \tilde{\omega})$ *there exists a square integrable solution of* (3.8). *Also the law uniqueness holds in the same sense as in Proposition* 1.

The above proposition can be proved by a method similar to Proposition 1. Only point one has to be careful is that the SDE (2.8) is now replaced by

$$X_n(t)=X(0)+\int_{S_t} b(X_n(h_n(s)),\ Y(h_n(s),\tilde{\omega}),\theta,R_n\sigma)dM \\ -c\int_0^t \{X_n(h_n(s))-\overline{Y(h_n(s))}\}ds,$$

where

$$R_n=R_n(s,\omega,\tilde{\omega}) \\ =R(X(s-,\omega),\ Y(s,\tilde{\omega}),X_n(h_n(s),\omega),\ Y(h_n(s),\tilde{\omega}))R(s,\omega,\tilde{\omega}).$$

Next we consider the SDE

$$(3.9)\quad X(t)=X(0)+\int_{S_t} b(X(s-),\ \tilde{X}(s,\tilde{\omega}),\theta,R(s,\omega,\tilde{\omega})\sigma)dM-c\int_0^t \{X(s)-\overline{X(s)}\}ds,$$

where $\tilde{X}(t,\tilde{\omega})$ satisfies the same conditions as stated in (2.14b). Then the following theorem can be proved in the same spirit as in Theorem 2.

THEOREM 4. *Under the condition* (II) *the probability measure on* W *induced by any square integrable solution of* (3.9) *is uniquely determined by the probability distribution* u_0 *of* $X(0,\omega)$.

REMARK 4. The martingale problems corresponding to the SDE's (2), (3), (2.14), (3.2), (3.3) and (3.9) have the same form and the probability distribution $u(t)$, at time t, of a solution to any one of these SDE's is a weak solution of the Boltzmann equation (1).

References

[1] Itô, K., On stochastic differential equations, Mem. Amer. Math. Soc. **4** (1951).
[2] McKean, H. P., A class of Markov processes associated with nonlinear parabolic equations, Proc. Nat. Acad. Sci. USA **56** (1966), 1907-1911.
[3] Murata, H., Propagation of chaos for Boltzmann-like equation of non-cutoff type in the plane, Hiroshima Math. J. **7** (1977), 479-515.
[4] Sznitman, A. S., Équations de type de Boltzmann, spatialement homogènes, Z. Wahrsch. Verw. Gebiete **66** (1984), 559-592.

[5] Tanaka, H., On Markov process corresponding to Boltzmann's equation of Maxwellian gas, Proc. 2nd Japan-USSR Symp. on Probab. Theory, Lecture Notes in Math. vol. 330, Springer-Verlag, Berlin-Heidelberg-New York, 1973, 478–489.
[6] Tanaka, H., Stochastic differential equation associated with the Boltzmann equation of Maxwellian molecules in several dimensions, Proc. of Conf. on Stochastic Analysis (edited by A. Friedman and M. Pinsky), Academic Press, New York-San Fancisco-London, 1978, 301-314.
[7] Tanaka, H., Probabilistic treatment of the Boltzmann equation of Maxwellian molecules, Z. Wahrsch. Verw. Gebiete **46** (1978), 67-105.

(Received September 19, 1986)

Department of Mathematics
Faculty of Science and Technology
Keio University
Yokohama
223 Japan

LIMIT THEOREM FOR ONE-DIMENSIONAL DIFFUSION PROCESS IN BROWNIAN ENVIRONMENT

Hiroshi TANAKA
Department of Mathematics. Faculty of Science and Technology
Keio University, Yokohama, 223 (Japan)

INTRODUCTION

Let $\mathbb{W}$ be the space of continuous functions $W: \mathbb{R} \to \mathbb{R}$ with $W(0) = 0$. In this paper an element of $\mathbb{W}$ is called an environment. Given an environment W, Brox[1] considered a diffusion process starting at 0 with generator

$$(1) \qquad \mathcal{L}_W = \frac{1}{2} e^{W(x)} \frac{d}{dx}\left(e^{-W(x)} \frac{d}{dx}\right) .$$

Such a diffusion, denoted by $X(t, W)$, is constructed from a one-dimensional Brownian motion $B(t)$ by a scale-change and a time-change. The probability measure governing $B(t)$ is denoted by P. We consider the Wiener measure Q on $\mathbb{W}$. Thus $\{W(x), x \geq 0, Q\}$ and $\{W(-x), x \geq 0, Q\}$ are independent one-dimensional Brownian motions starting at 0. We assume that $B(t)$ and $W(x)$ are independent so the full distribution governing $X(t, W)$ is $\mathcal{P} = P \otimes Q$. If $W(\cdot)$ were smooth, $X(t, W)$ would satisfy

$$X(t) = \text{a Brownian motion} - \frac{1}{2}\int_0^t W'(X(s))ds .$$

Although our $W(\cdot)$ is never smooth, the above remark will explain that $X(t, W)$ is regarded as a diffusion analogue of Sinai's random walk in a random environment([11]). $X(t, W)$ is called a diffusion process in a Brownian environment. The problem is to study the limiting behavior of $X(t, W)$ as $t \to \infty$. Brox[1] obtained the following result which is analogous to that of Sinai[11]: For any $\varepsilon > 0$

$$P\{(\log t)^{-2} X(t, W) \in U_\varepsilon(t, W)\} ,$$

which is regarded as a W-random variable, converges to 1 in probability as $t \to \infty$, where $U_\varepsilon(t, W)$ is the ε-neighborhood of $b_t(W)$ which is defined suitably in terms of "valleys" of the environment. The distribution of $b_t(W)$ is independent of t so the full distribution of $(\log t)^{-2} X(t, W)$ converges to that of $b_1(W)$. Kesten[7] obtained the exact form of the limit distribution. Kesten's result was then extended to the case of symmetric stable environments([12]).

The purpose of this paper is to elaborate Brox's result. We prove that, without scaling but only by centering, $X(t,\cdot)$ has a limit distribution as $t \to \infty$. To state the result more precisely, put $b(t, W) = (\log t)^2 b_t(W)$ and let $\tilde{Q}$ be the probability measure on $\mathbb{W}$ such that $\{W(x),\ x \geq 0,\ \tilde{Q}\}$ and $\{W(-x),\ x \geq 0,\ \tilde{Q}\}$ are independent Bessel processes of index 3 starting at 0. Note that $e^{-W} \in L^1(\mathbb{R})$ with $\tilde{Q}$-measure 1. Let Ω be the space of continuous paths $\omega: [0, \infty) \to \mathbb{R}$ and, for each W with $e^{-W} \in L^1(\mathbb{R})$, denote by $\tilde{P}_W$ the probability measure on Ω such that $\{\omega(t),\ t \geq 0,\ \tilde{P}_W\}$ is a diffusion process with generator (1) and with initial distribution

$$\tilde{\mu}_W(dx) = e^{-W(x)}dx \Big/ \int_{-\infty}^{\infty} e^{-W(y)}dy .$$

Finally put $\tilde{\mu} = \int \tilde{\mu}_W \tilde{Q}(dW)$ and $\tilde{P} = \int \tilde{P}_W \tilde{Q}(dW)$. Then our result is stated as follows.

__Theorem__. The process $\{X(t_0 + t, W) - b(t_0, W),\ t \geq 0,\ \mathcal{P}\}$ converges as $t_0 \to \infty$ to the stationary process $\{\omega(t),\ t \geq 0,\ \tilde{P}\}$ in the sense of weak convergence of probability measures on Ω. In particular the distribution of $X(t,\cdot) - b(t,\cdot)$ converges to $\tilde{\mu}$ as $t \to \infty$.

Similar results were also obtained by Golosov[2] for a reflecting random walk in random environment. Our method is on the extension line of Brox's and uses fine results on one-dimensional Brownian motion obtained by Lévy[B], Itô and McKean[4] and others.

§1. OUTLINE OF BROX'S METHOD

1.1. Let Ω_0 be the space of continuous paths $\omega: [0, \infty) \to \mathbb{R}$ with $\omega(0) = 0$ and let P be the Wiener measure on Ω_0. We write $B(t)$ for $\omega(t)$, the value of ω at time t. Thus $\{B(t),\ t \geq 0,\ P\}$ is a Brownian motion starting at 0. For a fixed $W \in \mathbb{W}$ we set

$$S(x) = \int_0^x e^{W(y)}dy ,$$

$S^{-1}(y)$ = the inverse function of $S(x)$,

$$A(s) = \int_0^s e^{-2W(S^{-1}(B(r)))}dr,$$

$A^{-1}(t)$ = the inverse function of $A(s)$.

Then $X(t, W) = S^{-1}(B(A^{-1}(t)))$ is a diffusion process with generator (1) starting at 0. If we set $(W^{x_0})(\cdot) = W(\cdot + x_0) - W(x_0)$, then

$X^{x_0}(t, W) = x_0 + X(t, W^{x_0})$ is a diffusion process with generator (1) starting at x_0.

Regarding W as a random element as well as ω, we have a process $X(t,\cdot)$ defined on the product probability space $\{\Omega \times \mathbb{W}, \mathcal{P} = P \otimes Q\}$. This full process is denoted by $\{X(t,\cdot)\}$ to distinguish it from the diffusion process $\{X(t, W), t \geq 0, P\}$ with a fixed W.

For $\lambda > 0$ and $W \in \mathbb{W}$ we define $W_\lambda \in \mathbb{W}$ by $W_\lambda(x) = \lambda^{-1}W(\lambda^2 x)$, $x \in \mathbb{R}$. The following scaling relation is important([1]): For any fixed $\lambda > 0$ and $W \in \mathbb{W}$

(1.1) $\quad \{X(t, \lambda W_\lambda), t \geq 0, P\} \overset{d}{=} \{\lambda^{-2}X(\lambda^4 t, W), t \geq 0, P\}$

where $\overset{d}{=}$ means the equality in distribution.

1.2. We give the definition of a valley. Let $W \in \mathbb{W}$. A part $\{W(x), a \leq x \leq c\}$ of W is called a *valley* of W if

(i) $a < c$,

(ii) there exists $b \in (a, c)$ such that

$W(a) > W(x) > W(b)$ for every $x \in (a, b)$,

$W(c) > W(x) > W(b)$ for every $x \in (b, c)$,

(iii) for the same b as above

$H_- \equiv \sup\{W(y) - W(x) : a \leq x < y \leq b\} < W(c) - W(b)$,

$H_+ \equiv \sup\{W(x) - W(y) : b \leq x < y \leq c\} < W(a) - W(b)$.

The value b in the above definition is particularly important and so, to stress b and also for simplicity, we write (a, b, c) instead of $\{W(x), a \leq x \leq c\}$. $A = H_+ \vee H_-$ is called the *inner directed ascent* (abbreviated to i.d.a.) and $D = (W(a) - W(b)) \wedge (W(c) - W(b))$ is the *depth* of the valley. When the letters A and D are used for notation they always mean the i.d.a. and the depth of (a,b,c). A valley (a,b,c) is said to contain 0 if $a < 0 < c$. It is easy to see that if (a, b, c) and (a', b', c') are valleys of W both containing 0 and satisfying $A < r < D$, $A' < r < D'$ (r is a positive constant), then $b = b'$. It is known that for any $r > 0$ and for any W in *some subset* $\hat{\mathbb{W}}$ $(\subset \mathbb{W})$ with Q-measure 1 there exists a valley (a, b, c) of W with $A < r < D$ and containing 0 (see[1]). We denote by $b(W)$ the unique b of such a valley (a, b, c) for $r = 1$.

To give another description of $b(W)$ we put

$$W^{\#}(x) = \begin{cases} W(x) - \min\limits_{[0,x]} W & \text{for } x \geq 0, \\ W(x) - \min\limits_{[x,0]} W & \text{for } x < 0, \end{cases}$$

$$d_+ = \inf\{x > 0 : W^{\#}(x) = 1\},$$
$$d_- = \sup\{x < 0 : W^{\#}(x) = 1\},$$
$$V_+ = \min_{[0,d_+]} W, \quad V_- = \min_{[d_-,0]} W,$$

and define b_+ and b_- by $W(b_\pm) = V_\pm$ (such $b_\pm$ are uniquely determined with Q-measure 1). We also set

$$M_+ = \max_{[0,b_+]} W, \quad M_- = \max_{[b_-,0]} W.$$

Then another description of $b(W)$ is given as follows(see[7]):

$$(1.2) \qquad b(W) = \begin{cases} b_+ & \text{if } M_+ \vee (V_+ + 1) < M_- \vee (V_- + 1), \\ b_- & \text{if } M_+ \vee (V_+ + 1) > M_- \vee (V_- + 1). \end{cases}$$

Moreover, if we define $a(W)$ and $c(W)$ by

$a(W)$ = the infimum of the set of a's ($a < b(W)$) such that
$$\begin{cases} W(a) > W(x) > W(b(W)) & \text{for every } x \in (a, b(W)), \\ \sup\{W(y) - W(x) : a \leq x < y \leq b(W)\} < 1, \end{cases}$$

$c(W)$ = the supremun of the set of c's ($c > b(W)$) such that
$$\begin{cases} W(c) > W(x) > W(b(W)) & \text{for every } x \in (b(W), c), \\ \sup\{W(x) - W(y) : b(W) \leq x < y \leq c\} < 1. \end{cases}$$

Then $(a(W), b(W), c(W))$ is the maximum valley of W containing 0 and satisfying $A(W) < 1 < D(W)$, and is called the *standard valley* of W. Note that $a(W)$, $b(W)$, etc., are Borel functions on $\hat{W}$.

1.3. Let (a, b, c) be a valley of W and, for $\lambda > 0$, let T_λ^x be the exit time from (a, c) for the diffusion process $X^x(t, \lambda W)$. The following lemma is due to Brox[1].

<u>*Lemma 1*</u> *([1]). For any $\delta > 0$ and a closed interval $I \subset (a, c)$*

$$(1.3) \qquad \lim_{\lambda\to\infty} \inf_{x\in I} P\{e^{\lambda(D-\delta)} < T_\lambda^x < e^{\lambda(D+\delta)}\} = 1 .$$

1.4. Brox([1]) employs a coupling technique. To explain it we had better adopt the path space representation of $X(t, \lambda W)$. So let $\Omega = C([0, \infty) \to \mathbb{R})$ and denote by $P_{\lambda W}$ the probability measure on Ω induced by the diffusion process $X(t, \lambda W)$. Moreover, we use the following notation. For an arbitrary interval $[a, c]$ and an environment W we denote by $P^\lambda_{W[a,c]}$ the probability measure on the path space $\Omega_{[a,c]} = C([0, \infty) \to [a, c])$ induced by the diffusion process on $[a, c]$ with (local) generator $\mathcal{L}_{\lambda W}$, with reflecting barriers at a and c and with initial distribution

$$\mu^{\lambda}_{W[a,c]}(dx) = e^{-\lambda W(x)}dx \Big/ \int_a^c e^{-\lambda W(y)}dy \ .$$

This reflecting diffusion is stationary since $\mu^{\lambda}_{W[a,c]}$ is its invariant measure. In particular, in case (a, b, c) is the standard valley of W, $P^{\lambda}_{W[a,c]}$, $\mu^{\lambda}_{W[a,c]}$ and $\Omega_{[a,c]}$ are abbreviated to P^{λ}_{W}, μ^{λ}_{W} and $\hat{\Omega}$, respectively.

We now assume that (a, b, c) is the standard valley of W. Let ω and $\hat{\omega}$ stand for generic elements of Ω and $\hat{\Omega}$ with values $\omega(t)$ and $\hat{\omega}(t)$ at time t, respectively, and consider two processes $\{\omega(t),\ t \geq 0\}$ and $\{\hat{\omega}(t),\ t \geq 0\}$ defined on the product probability space $\{\Omega^{\lambda} \times \hat{\Omega},\ \mathbb{P}^{\lambda}_{W}\}$ where $\mathbb{P}^{\lambda}_{W} = P_{\lambda W} \otimes P^{\lambda}_{W}$. Thus the two processes are independent. Put

$$R = \inf\{t \geq 0 : \omega(t) = \hat{\omega}(t)\} \ ,$$
$$T_R = \inf\{t \geq R : \omega(t) \notin (a,\ c)\} \ ,$$
$$\hat{T}_R = \inf\{t \geq R : \hat{\omega}(t) \notin (a,\ c)\} \ .$$

Notice that these are random variables defined on $\{\Omega \times \hat{\Omega},\ \mathbb{P}^{\lambda}_{W}\}$. If we define a process $\{\omega'(t),\ t \geq 0\}$ by

$$\omega'(t) = \begin{cases} \omega(t) & \text{for } 0 \leq t \leq R\ , \\ \hat{\omega}(t) & \text{for } t > R\ , \end{cases}$$

then

$$(1.4)\qquad \{\omega(t),\ 0 \leq t \leq T_R,\ \mathbb{P}^{\lambda}_{W}\} \overset{d}{=} \{\omega'(t),\ 0 \leq t \leq \hat{T}_R,\ \mathbb{P}^{\lambda}_{W}\} \ .$$

The following lemma is also due to Brox[1]; the equality is a consequence of (1.4).

Lemma 2 ([1]). For any r_1 and r_2 such that $A < r_1 < r_2 < D$

$$(1.5)\qquad \mathbb{P}^{\lambda}_{W}\{R < e^{\lambda r_1} < e^{\lambda r_2} < T_R\}$$
$$= \mathbb{P}^{\lambda}_{W}\{R < e^{\lambda r_1} < e^{\lambda r_2} < \hat{T}_R\} \to 1, \quad \lambda \to \infty \ .$$

Using Lemma 2 and the scaling relation (1.1), Brox([1]) obtained his main results: For any $\varepsilon > 0$

$$(1.6)\qquad P\{|\lambda^{-2}X(e^{\lambda},\ W) - b(W_{\lambda})| > \varepsilon\} \to 0$$

in probability with respect to Q as $\lambda \to \infty$.

By the same argument as Brox's we can obtain a refinement of his result as will be discussed in the next subsection.

1.5. We keep the notation of 1.4 and, in addition, we denote by θ_t (resp. $\hat{\theta}_t$) the shift on Ω (resp. $\hat{\Omega}$) defined by $(\theta_t\omega)(\cdot) = \omega(t + \cdot)$ (resp. $(\hat{\theta}_t\hat{\omega})(\cdot) = \hat{\omega}(t + \cdot)$). For $\lambda > 0$, γ_λ denotes the map: $\Omega \to \Omega$ defined by $(\gamma_\lambda\omega)(t) = \lambda^2\omega(\lambda^{-4}t)$, $t \geq 0$. $\mathcal{B}_t$ denotes the σ-field on Ω generated by the sets of the form $\{\omega: \omega(s) \leq x\}$, $0 \leq s \leq t$, $x \in \mathbb{R}$, and $\mathcal{B} = \bigvee \mathcal{B}_t$. For $\omega \in \Omega$ and $x \in \mathbb{R}$, $\omega - x$ denotes the path whose value at time t is $\omega(t) - x$; $\hat{\omega} - x$ (for $\hat{\omega} \in \Omega$) also denotes a similar path. The following notational convention is used:

$$(1.7a) \qquad P^\lambda_{W[a,c]}\{\Gamma\} = P^\lambda_{W[a,c]}\{\Gamma \cap \Omega_{[a,c]}\}, \quad \Gamma \in \mathcal{B},$$

$$(1.7b) \qquad P^\lambda_W\{\hat{\omega} \in \Gamma\} = P^\lambda_W\{\Gamma \cap \hat{\Omega}\}, \quad \Gamma \in \mathcal{B}.$$

Note that the right hand sides of the above make sense since both $\Omega_{[a,c]}$ and $\hat{\Omega}$ are measurable subsets of Ω.

For any family $\{r(\lambda)\}$ such that $r(\lambda) \to 1$ $(\lambda \to \infty)$ Lemma 2 implies

$$(1.8) \qquad \varepsilon_\lambda(W) \equiv 1 - \mathbb{P}^\lambda_W\{R < s(\lambda) < s(\lambda) + t(\lambda) < T_R\} \to 0$$

as $\lambda \to \infty$ for any $W \in \tilde{W}$ and the same is true with T_R replaced by $\hat{T}_R$, where

$$(1.9) \qquad s(\lambda) = \lambda^{-4}e^\lambda, \quad t(\lambda) = \lambda^{-4}e^{\lambda r(\lambda)}.$$

We are now in position to state

<u>*Refinement of Brox's result.*</u> *For* $\{r(\lambda)\}$ *as above and for any* $\Gamma_\lambda \in \mathcal{B}_{u(\lambda)}$, $u(\lambda) = e^{\lambda r(\lambda)}$, $\lambda > 0$, *we have*

$$(1.10) \qquad P_W\{\theta_{\exp\lambda}\omega - \lambda^2 b(W_\lambda) \in \Gamma_\lambda\}$$
$$\stackrel{d}{=} P^\lambda_W\{\hat{\omega} - b(W) \in \gamma_\lambda^{-1}(\Gamma_\lambda)\}^{*)} + \varepsilon_\lambda(W, \Gamma_\lambda),$$

where $b(\cdot)$ *is defined by* (1.2) *and* $\varepsilon_\lambda(\cdot, \Gamma_\lambda)$ *is a suitable random variable defined on* $(\tilde{W}, Q)$ *satisfying*

$$(1.11) \qquad |\varepsilon_\lambda(W, \Gamma_\lambda)| \leq \varepsilon_\lambda(W).$$

The proof is as follows. Since the scaling relation (1.1) implies

---

*) The convention (1.7b) is used.

$$\{(\theta_{\exp\lambda}\omega)(t),\ t \geq 0,\ P_W\} \overset{d}{=} \{(\gamma_\lambda \theta_{s(\lambda)}\omega)(t),\ t \geq 0,\ P_{\lambda W_\lambda}\}\ ,$$

using the notation (1.9) we have

$$\begin{aligned}
&P_W\{\theta_{\exp\lambda}\omega - \lambda^2 b(W_\lambda) \in \Gamma_\lambda\} \\
&= P_{\lambda W_\lambda}\{\gamma_\lambda \theta_{s(\lambda)}\omega - \lambda^2 b(W_\lambda) \in \Gamma_\lambda\} \\
&\overset{d}{=} P_{\lambda W}\{\gamma_\lambda \theta_{s(\lambda)}\omega - \lambda^2 b(W) \in \Gamma_\lambda\} \qquad (\text{since } W_\lambda \overset{d}{=} W) \\
&= \mathbb{P}_W^\lambda\{R < s(\lambda),\ \theta_{s(\lambda)}\omega - b(W) \in \gamma_\lambda^{-1}(\Gamma_\lambda),\ s(\lambda)+t(\lambda) < T_R\} + \varepsilon_\lambda^0 \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad (\text{by } (1.8)) \\
&= \mathbb{P}_W^\lambda\{R < s(\lambda),\ \hat\theta_{s(\lambda)}\omega - b(W) \in \gamma_\lambda^{-1}(\Gamma_\lambda),\ s(\lambda)+t(\lambda) < \hat T_R\} + \varepsilon_\lambda^0 \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad (\text{by } (1.4)) \\
&= P_W^\lambda\{\hat\theta_{s(\lambda)}\ \hat\omega - b(W) \in \gamma_\lambda^{-1}(\Gamma_\lambda)\} + \varepsilon_\lambda \qquad (\text{by } (1.8) \text{ for } \hat T_R) \\
&= P_W^\lambda\{\hat\omega - b(W) \in \gamma_\lambda^{-1}(\Gamma_\lambda)\} + \varepsilon_\lambda,
\end{aligned}$$

where $\varepsilon_\lambda^0 = \varepsilon_\lambda^0(W, \Gamma_\lambda)$ and $\varepsilon_\lambda = \varepsilon_\lambda(W, \Gamma_\lambda)$ are suitable random variables satisfying (1.11).

Before ending this section we state one more lemma.

Lemma 3. *Let (a, b, c) be the standard valley of W $(\in \tilde{W})$. If $\rho(\lambda)$, $\lambda > 0$, satisfies*

(1.12) $\quad \rho(\lambda) \geq 0$ *and* $\rho(\lambda) = o(e^{\varepsilon\lambda})$ *as* $\lambda \to \infty$ *for* $\forall\varepsilon > 0$,

then for any Γ_λ, $\lambda > 0$, *satisfying*

(1.13) $\quad \Gamma_\lambda \in \mathcal{B}_{\rho(\lambda)}$, $\lambda > 0$,

and for any a' *and* c' *with* $a < a' < b < c' < c$ *we have*

(1.14) $\quad P_W^\lambda\{\hat\omega - b \in \Gamma_\lambda\} = P^\lambda_{W^b[a'-b,c'-b]}\{\Gamma_\lambda\} + o(1)$, $\lambda \to \infty$,

where $o(1)$ *is uniform with respect to the choice of* Γ_λ *under the condition (1.13) and* $W^b(\cdot) = W(\cdot + b) - W(b)$.

Proof. Since for any $\delta > 0$ both $\mu_W^\lambda\{(b - \delta, b - \delta)\}$ and $\mu^\lambda_{W^b[a'-b,c'-b]}\{(-\delta,-\delta)\}$ tend to 1 as $\lambda \to \infty$, it follows from Lemma 1 and (1.12) that

(1.15) $\quad P_W^\lambda\{\hat T > \rho(\lambda)\} \to 1,\ P^\lambda_{W^b[a'-b,c'-b]}\{T' > \rho(\lambda)\} \to 1 \quad (\lambda \to \infty)$, where $\hat{T}$ and T' are the exit times of (a', c') and $(a'-b, c'-b)$, respectively, for the processes under consideration. Moreover, we see that

(1.16a) $\mu_W^\lambda(dx) = \kappa(\lambda)\mu_{W[a',c']}^\lambda(dx)$ on $[a', c']$,

(1.16b) $\kappa(\lambda) = \int_{a'}^{c'} e^{-\lambda W(x)}dx \Big/ \int_a^c e^{-\lambda W(x)}dx \to 1, \quad \lambda \to \infty$.

Therefore, we have as $\lambda \to \infty$

$$
\begin{aligned}
& P_W^\lambda\{\hat{\omega} - b \in \Gamma_\lambda\} \\
&= P_W^\lambda\{\hat{\omega} - b \in \Gamma_\lambda, \hat{T} > \rho(\lambda)\} + o(1) \qquad \text{(by (1.15))} \\
&= \kappa(\lambda) P^\lambda_{W^b[a'-b,c'-b]}[\Gamma_\lambda \cap \{T' > \rho(\lambda)\}] + o(1) \qquad \text{(by 1.16a))} \\
&= P^\lambda_{W^b[a'-b,c'-b]}[\Gamma_\lambda \cap \{T' > \rho(\lambda)\}] + o(1) \qquad \text{(by 1.16b))} \\
&= P^\lambda_{W^b[a'-b,c'-b]}\{\Gamma_\lambda\} + o(1) \qquad \text{(by (1.15))} ,
\end{aligned}
$$

completing the proof of the lemma.

§2. THE LAW OF THE STANDARD VALLEY.

Recalling the notation of 1.2 we put

$$W_+ = \{W(b_+ + t) - W(b_+), -b_+ \le t \le d_+ - b_+, Q\} ,$$
$$W_- = \{W(b_- - t) - W(b_-), b_- \le t \le -(d_- - b_-), Q\} .$$

On a suitable probability space we consider a process $\{x_+(t), t \in \mathbb{R}\}$ such that $\{x_+(t), t \ge 0\}$ and $\{x_+(-t), t \ge 0\}$ are independent reflecting Brownian motions on $[0, \infty)$ starting at 0 (abbreviation: RBM^0). Let $\{\ell_+(t), t \in \mathbb{R}\}$ be the local timeat 0 of $x_+(t)$, that is,

$$\ell_+(t) = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_I \mathbf{1}_{[0,\varepsilon]}(x_+(s))ds ,$$

where $I = [0, t]$ or $[t, 0]$ according as $t \ge 0$ or $t < 0$. Also put $\beta_+(t) = x_+(t) + \ell_+(t)$. Then by Pitman's theorem([9]) $\{\beta_+(t), t \ge 0\}$ is a Bessel process of index 3 starting at 0 (abbreviation: $BES^0(3)$). We put

(2.1a) σ_+ = the smallest zero of $x_+(t)$ in $(z, 0]$ where z is the maximum of $t < 0$ with $x_+(t) = 1$,

(2.1b) $\tau_+ = \min\{t > 0 : \beta_+(t) = 1\}$.

Proposition. W_- *and* W_+ *are independent and*

(2.2) $$W_- \stackrel{d}{=} W_+ \stackrel{d}{=} \{\beta_+(t),\ \sigma_+ \le t \le \tau_+\} .$$

Proof. The independence of W_- and W_+ and the first law equality in (2.2) are obvious. For the proof of the second law equality it is convenient to use the construction of an equivalent of W_+ by means of the excursions of a RBM^0 ([4][5][3]). Denote by $\mathcal{W}^+$ the space of $w:[0,\infty) \to \mathbb{R}^+$ satisfying

(i) $w(t) > 0$ for $0 < t < \zeta(w) = \min\{s > 0 : w(s) = 0\}$,

(ii) $w(0) = w(t) = 0$ for $t \ge \zeta(w)$.

We consider a σ-finite measure n^+ on $\mathcal{W}^+$ defined by

$$n^+\left[\{w(t_1) \in A_1,\ w(t_2) \in A_2, \cdots, w(t_n) \in A_n\}\right]$$

$$= \int_{A_1} K^+(t_1, x_1)dx_1 \int_{A_2} p^0(t_2-t_1, x_1, x_2)dx_2 \int_{A_3} \cdots$$

$$\cdots \int_{A_n} p^0(t_n - t_{n-1}, x_{n-1}, x_n)dx_n$$

where $0 < t_1 < t_2 < \cdots < t_n$, $A_i \in \mathcal{B}(\mathbb{R}^+)$, $1 \le i \le n$, and

$$K^+(t, x) = \sqrt{\frac{2}{\pi t^3}}\, e^{-x^2/2t} , \quad t > 0, \quad x \in \mathbb{R}^+,$$

$$p^0(t,x,y) = \frac{1}{\sqrt{2\pi t}}\{e^{-(x-y)^2/2t} - e^{-(x+y)^2/2t}\} ,$$

$$t > 0, \quad x, y \in \mathbb{R}^+.$$

Let $p(t)$ be a stationary Poisson point process on $\mathcal{W}^+$ with characteristic measure n^+ and set

$$a(t) = \sum_{0<s\le t} \zeta(p(s)) ,$$

$$\ell(t) = \text{the inverse function of } a(\cdot) .$$

We define $x(t)$ by

$$x(t) = \begin{cases} p(s)(t - a(s-)) & \text{if } a(s-) \le t \le a(s) \\ 0 & \text{if } a(s-) = t = a(s) , \end{cases}$$

where $s = \ell(t)$. Then $x(t)$ is a RBM^0 and $\ell(t)$ is its local time at 0 ([3]). For $w \in \mathcal{W}^+$ we denote by $h(w)$ the maximum value of $w(t)$ and put

$$\xi = \min\{s > 0 : h(p(s)) > 1\} ,$$

$$\tau = \min\{t > 0 : p(\xi)(t) = 1\} .$$

Note that $\xi < \infty$ a.s. because $n^+[\{h(w) > 1\}] = 1 < \infty$. We also put

$$W_+^- = \{W(b_+ + t) - W(b_+),\ -b_+ \le t \le 0,\ Q\} ,$$

$$W_+^+ = \{W(b_+ + t) - W(b_+),\ 0 \le t \le d_+ - b_+,\ Q\} ,$$

$$W_+^0 = \{W(t),\ 0 \le t \le b_+,\ Q\} .$$

Since $\{W^\#(t),\ t \ge 0,\ Q\}$ is a RBM^0 and $-\min\{W(s): 0 \le s \le t\}$ is its local time at 0, we see that the joint distribution of W_+^0 and W_+^+ is the same as the joint distribution of the following processes (2.3) and (2.4):

(2.3) $\{x(t) - \ell(t),\ 0 \le t \le a(\xi-)\}$,

(2.4) $\{p(\xi)(t),\ 0 \le t \le \tau\}$.

On the other hand, $\{p(s),\ 0 \le s < \xi\}$ and $\{p(\xi)(t),\ 0 \le t \le \tau\}$ are independent and hence the processes (2.3) and (2.4) are also independent. Therefore, W_+^+ is independent of W_+^0 and consequently of W_+^-. Moreover, since $\{p(\xi)(t),\ 0 \le t \le \tau\}$is a part of a $BES^0(3)$ ([13], see also [10]), it follows that

$$W_+^+ \overset{d}{=} \{\beta_+(t),\ 0 \le t \le \tau_+\} .$$

It remains to prove

(2.5) $W_+^- \overset{d}{=} \{\beta_+(t),\ \sigma_+ \le t \le 0\}$.

To prove this we must consider the time reversal of (2.3), or more precisely, of

(2.6) $\{x(t) - \ell(t) + \xi,\ \ 0 \le t \le a(\xi-)\}$.

Put

$$\mathcal{W}_0^+ = \{w \in \mathcal{W}^+ : h(w) < 1\} ,$$

$$\mathcal{W}_1^+ = \{w \in \mathcal{W}^+ : h(w) \ge 1\} ,$$

and define two point processes p_0 and p_1 by

$$p_i(t) = p(t) \qquad \text{for } t \in D_{p_i},\quad i = 0, 1 ,$$

where D_{p_i} (the domain of definition of p_i) is given by

$$D_{p_i} = \{t \in (0, \infty) : p(t) \in \mathcal{W}_i^+\},\quad i = 0, 1 .$$

Then p_0 and p_1 are independent stationary Poisson point processes on $\mathcal{W}_0^+$ and $\mathcal{W}_1^+$ with characteristic measures $n_i^+ =$ the restriction of n^+ to $\mathcal{W}_i^+$, i = 0, 1, respectively. For $w \in \mathcal{W}_0^+$ we define $\hat{w} \in \mathcal{W}_0^+$ by

$$\hat{w}(s) = \begin{cases} w(\zeta(w) - s) & \text{for } 0 \leq s \leq \zeta(w) \\ 0 & \text{for } s > \zeta(w) . \end{cases}$$

Then the measure n_0^+ is invariant under the map: $w \to \hat{w}$. We define $\tilde{p}_0$ by

$$\tilde{p}_0(t) = \begin{cases} \widehat{p_0(\xi - t)} & \text{for } t < \xi \\ p_0(t) & \text{for } t \geq \xi . \end{cases}$$

It can be proved that $\tilde{p}_0$ is again a stationary Poisson point process on $\mathcal{W}_0^+$ with characteristic measure n_0^+ and is independent of p_1. Therefore, the point process $\tilde{p}$ defined by

$$\tilde{p}(t) = \begin{cases} \tilde{p}_0(t) & \text{for } t \in D_{\tilde{p}_0} \\ p_1(t) & \text{for } t \in D_{p_1} \end{cases}$$

is equivalent in law to p. Therefore, if we set

$$\tilde{a}(t) = \sum_{0<s\leq t} \zeta(\tilde{p}(t)) ,$$

$$\tilde{\ell}(t) = \text{the inverse function of } \tilde{a}(\cdot) ,$$

and if we define $\tilde{x}(t)$ by

$$\tilde{x}(t) = \begin{cases} \tilde{p}(s)(t - \tilde{a}(s-)) & \text{if } \tilde{a}(s-) \leq t \leq \tilde{a}(s) \\ 0 & \text{if } \tilde{a}(s-) = t = \tilde{a}(s) \end{cases}$$

where $s = \tilde{\ell}(t)$, then $\tilde{x}(t)$ is a RBM^0 and $\tilde{\ell}(t)$ is its local time at 0. Also we have $\tilde{a}(\xi-) = a(\xi-)$. Moreover, it is seen that (drawing a picture is helpful)

$$\tilde{x}(t) = x(a(\xi-) - t) , \qquad 0 \leq t \leq a(\xi-) ,$$

$$\tilde{\ell}(t) = \ell(a(\xi-)) - \ell(a(\xi-) - t)$$

$$= \xi - \ell(a(\xi-) - t) , \qquad 0 \leq t \leq a(\xi-) .$$

Therefore, the time reversal of (2.6) is

$$\text{(2.7)} \qquad \{x(a(\xi-) - t) - \ell(a(\xi-) - t) + \xi, \ 0 \leq t \leq a(\xi-)\}$$

$$\stackrel{d}{=} \{\tilde{x}(t) + \tilde{\ell}(t), \ 0 \leq t \leq \tilde{a}(\xi-)\} .$$

Since W_+^- is equivalent in law to

$$\{x(a(\xi-) + t) - \ell(a(\xi-) + t) + \xi, \ -a(\xi-) \leq t \leq 0\} ,$$

(2.7) implies (2.5) as was to be proved.

§3. PROOF OF THE THEOREM

We are going to prove the theorem announced in the introduction with

$$b(t,W) = (\log t)^2 b(W_{\log t}), \quad t > 1 . \tag{3.1}$$

Let $\mathbb{P}$ be the probability measure on $\Omega \times \mathbb{W}$ defined by $\mathbb{P}(d\omega dW) = P_W(d\omega)Q(dW)$. Then the theorem is rephased as follows: The process

$$\{\omega(e^\lambda + t) - \lambda^2 b(W_\lambda), \quad t \geq 0, \quad \mathbb{P}\}$$

converges in law to the process $\{\omega(t), t \geq 0, \bar{P}\}$ as $\lambda \to \infty$.

In addition to the process $x_+(t)$ of §2, we need another process $\{x_-(t), t \in \mathbb{R}\}$ which is equivalent in law to $\{x_+(t), t \in \mathbb{R}\}$. We assume that $x_+(t)$ and $x_-(t)$ are defined on a common probability space $\{\bar{\Omega}, \bar{P}\}$ and that they are independent. Denote by $\ell_-(t)$ the local time at 0 of $x_-(t)$, put $\beta_-(t) = x_-(t) + \ell_-(t)$ and define σ_- and τ_- in a way similar to (2.1). The expectation with respect to $\bar{P}$ is denote by $\bar{E}$.

Let $\rho(\lambda)$, $\lambda > 0$, be a given function satisfying the condition (1.12) and let $\Gamma_\lambda \in \mathcal{B}_{\rho(\lambda)}$, $\lambda > 0$. Then the condition for Γ_λ in the refinement of Brox's result is automatically satisfied. Therefore, by (1.10) and (1.2) we have

$$\begin{aligned} &E_Q[P_W\{\theta_{\exp\lambda}\omega - \lambda^2 b(W_\lambda) \in \Gamma_\lambda\}] \\ &= E_Q[P_W^\lambda\{\hat{\omega} - b(W) \in \gamma_\lambda^{-1}(\Gamma_\lambda)\}] + o(1) \\ &= I_\lambda + II_\lambda + o(1) , \end{aligned} \tag{3.2}$$

where

$$I_\lambda = E_Q[P_W^\lambda\{\hat{\omega} - b_+ \in \gamma_\lambda^{-1}(\Gamma_\lambda)\}; b = b_+]$$

$$II_\lambda = E_Q[P_W^\lambda\{\hat{\omega} - b_- \in \gamma_\lambda^{-1}(\Gamma_\lambda)\}; b = b_-] ,$$

and E_Q denotes the expectation with respect to Q while $E_Q\{—; b = b_+\}$ denotes the integral over the set $\{b = b_+\}$. It is to be noted that in the above (and also in what follows) $o(1)$ means a term which tends to 0 as $\lambda \to \infty$ uniformly with respect to the choice of $\{\Gamma_\lambda\}$ so far as it satisfies the condition (1.13). We are going to use (2.2) to compute I_λ and II_λ.

We put

$$\bar{b}_+ = -\sigma_+, \quad \bar{b}_- = \sigma_- ,$$

16B

$$\bar{d}_+ = \tau_+ - \sigma_+, \quad \bar{d}_- = -(\tau_- - \sigma_-) ,$$

$$\bar{M}_+ = \max_{[\sigma_+,0]} \beta_+ - \beta_+(\sigma_+), \quad \bar{M}_- = \max_{[\sigma_-,0]} \beta_- - \beta_-(\sigma_-) ,$$

$$\bar{V}_+ = -\beta_+(\sigma_+), \quad \bar{V}_- = -\beta_-(\sigma_-) ,$$

$$\beta(t) = \begin{cases} \beta_+(\sigma_+ + t) - \beta_+(\sigma_+) & \text{for } t \geq 0 , \\ \beta_-(\sigma_- - t) - \beta_-(\sigma_-) & \text{for } t < 0 , \end{cases}$$

Then the proposition of §2 implies

$$(3.3) \qquad \{W(t),\ d_- \leq t \leq d_+\} \overset{d}{=} \{\beta(t),\ \bar{d}_- \leq t \leq \bar{d}_+\} .$$

As in (1.2) we define $\bar{b}$ by

$$\bar{b} = \begin{cases} \bar{b}_+ & \text{if } \bar{M}_+ \vee (\bar{V}_+ + 1) < \bar{M}_- \vee (\bar{V}_- + 1) , \\ \bar{b}_- & \text{if } \bar{M}_+ \vee (\bar{V}_+ + 1) > \bar{M}_- \vee (\bar{V}_- + 1) . \end{cases}$$

Using Lemma 3 and then (3.3) we have

$$(3.4) \qquad I_\lambda = E_Q[P^\lambda_{W^{b+}[-b_+,\ d_+-b_+]}\{\gamma_\lambda^{-1}(\Gamma_\lambda)\};\ b = b_+] + o(1)$$

$$= \bar{E}[P^\lambda_{\beta_+[\sigma_+,\ \tau_+]}\{\gamma_\lambda^{-1}(\Gamma_\lambda)\};\ b = b_+] + o(1) ,$$

where $P^\lambda_{\beta_+[\cdot,\cdot]}$ is the probability measure defined in a way similar to $P^\lambda_{W[\cdot,\cdot]}$.

For $\varepsilon > 0$ put

$$\sigma_\varepsilon = \max\{t < 0 : x_+(t) = 0 \text{ and } t < {}^\exists s < 0 \text{ s.t. } x_+(s) = \varepsilon\},$$

$$P^\lambda_{\beta_+,\varepsilon} = P^\lambda_{\beta_+[\sigma_\varepsilon,\tau_+]} .$$

Also let $\tilde{\mu}_{\lambda\beta_+}$ be the probability measure on $\mathbb{R}$:

$$\tilde{\mu}_{\lambda\beta_+}(dx) = e^{-\lambda\beta_+(x)}dx \Big/ \int_{-\infty}^{\infty} e^{-\lambda\beta_+(t)}dt ,$$

and let $\bar{P}_{\lambda\beta_+}$ be the probability measure on Ω defined in a way similar to $\hat{P}_{\lambda W}$ (see the introduction). Then the scaling relation

$$(3.5) \qquad \{\lambda^{-1}\beta_+(\lambda^2 t),\ t \in \mathbb{R}\} \overset{d}{=} \{\beta_+(t),\ t \in \mathbb{R}\}$$

implies that for any $\delta > 0$

$$\tilde{\mu}_{\lambda\beta_+}\{(-\delta,\delta)\} \overset{d}{=} \int_{-\lambda^2\delta}^{\lambda^2\delta} e^{-\beta_+(x)}dx \Big/ \int_{-\infty}^{\infty} e^{-\beta_+(t)}dt \to 1, \quad \lambda \to \infty ,$$

and hence $E[\tilde{\mu}_{\lambda\beta_+}\{(-\delta,\delta)\}] \to 1$. Similarly $E[\mu^{\lambda}_{\beta_+[\sigma_+,\tau_+]}\{(-\delta,\delta)\}] \to 1$ and $E[\mu^{\lambda}_{\beta_+[\sigma_\varepsilon,\tau_+]}\{(-\delta,\delta)\}] \to 1$ as $\lambda \to \infty$. Therefore, just as in the case of Lemma 3, we can prove the following: Let Γ_λ, $\lambda > 0$, be the same as before. Then

$$(3.6) \qquad \bar{E}[\tilde{P}_{\lambda\beta_+}\{\gamma_\lambda^{-1}(\Gamma_\lambda)\}] = \bar{E}[P^{\lambda}_{\beta_+[\sigma_+,\tau_+]}\{\gamma_\lambda^{-1}(\Gamma_\lambda)\}] + o(1)$$
$$= \bar{E}[P^{\lambda}_{\beta_+,\varepsilon}\{\gamma_\lambda^{-1}(\Gamma_\lambda)\}] + o(1), \quad \lambda \to \infty .$$

From (3.4) and (3.6) we have for any $\varepsilon > 0$

$$(3.7) \qquad I_\lambda = \bar{E}[P^{\lambda}_{\beta_+,\varepsilon}\{\gamma_\lambda^{-1}(\Gamma_\lambda)\};\ \bar{b} = \bar{b}_+] + o(1), \quad \lambda \to \infty ,$$

and a similar formula for II_λ.

We put

$\tilde{x}_+(t) = x_+(\sigma_\varepsilon + t) - x_+(\sigma_\varepsilon)$,

$\tilde{\beta}_+(t) = \beta_+(\sigma_\varepsilon + t) - \beta_+(\sigma_\varepsilon)$,

$\tilde{\sigma}_+$ = the smallest zero of $\tilde{x}_+(t)$ in $(z, 0]$ where z is the maximum of $t < 0$ such that $\tilde{x}_+(t) = 1$,

$\tilde{M}_+ = \max\limits_{[\tilde{\sigma}_+,0]} \tilde{\beta}_+ - \tilde{\beta}_+(\tilde{\sigma}_+)$,

$m_+ = \max\limits_{[\sigma_+,0]} \beta_+, \quad m_\varepsilon = \max\limits_{[\sigma_\varepsilon,0]} \beta_+.$

Then $m_\varepsilon < m_+$ for all sufficiently small $\varepsilon > 0$ ($\bar{P}$-a.s.) and

$$m_\varepsilon < m_+ \Longrightarrow \sigma_+ < \sigma_\varepsilon$$
$$\Longrightarrow \beta_+(\sigma_+) = \tilde{\beta}_+(\tilde{\sigma}_+) + \beta_+(\sigma_\varepsilon), \quad \bar{M}_+ = \tilde{M}_+.$$

Therefore

$$\{\bar{b} = \bar{b}_+\} \cap \{m_\varepsilon < m_+\}$$
$$= \{\tilde{M}_+ \vee (1 - \tilde{\beta}_+(\tilde{\sigma}_+) - \beta_+(\sigma_\varepsilon)) < \bar{M}_- \vee (1 - \beta_-(\sigma_-))\} \cap \{m_\varepsilon < m_+\} ,$$

and hence

$$(3.8) \qquad \bar{E}[P^{\lambda}_{\beta_+,\varepsilon}\{\gamma_\lambda^{-1}(\Gamma_\lambda)\};\ \bar{b} = \bar{b}_+]$$
$$= \bar{E}[P^{\lambda}_{\beta_+,\varepsilon}\{\gamma_\lambda^{-1}(\Gamma_\lambda)\};\ \tilde{M}_+ \vee (1-\tilde{\beta}_+(\tilde{\sigma}_+)-\beta_+(\sigma_\varepsilon)) < \bar{M}_- \vee (1-\beta_-(\sigma_-))]$$
$$+ \Delta(\varepsilon,\lambda) ,$$

where $\Delta(\varepsilon,\lambda)$ tends to 0 as $\varepsilon\downarrow 0$ uniformly in λ. Since $P^{\lambda}_{\beta_+,\varepsilon}\{\gamma_\lambda^{-1}(\Gamma_\lambda)\}$ is a measurable function of $\{x_+(t),\ \sigma_\varepsilon\le t\le\tau_+\}$ and since the process $\{\tilde\beta_+(t),\ \tilde\sigma_+\le t\le 0\}$ conditioned by $\{x_+(t),\ \sigma_\varepsilon\le t\le\tau_+\}$ is equivalent in law to $\{\beta_+(t),\ \sigma_+\le t\le 0\}$, by using the strong Markov property of $x_+(t)$ we see that the right hand side of (3.8) equals

$$(3.9)\qquad \bar E[P^{\lambda}_{\beta_+,\varepsilon}\{\gamma_\lambda^{-1}(\Gamma_\lambda)\}\Psi(\beta_+(\sigma_\varepsilon))] + \Delta(\varepsilon,\lambda)\ ,$$

where $\psi(x) = \bar P\{\bar M_+^{\vee}(1-\beta_+(\sigma_+)-x) < \bar M_-^{\vee}(1-\beta_-(\sigma_-))\}$. Since $\beta_+(\sigma_\varepsilon)\to 0$ as $\varepsilon\downarrow 0$ and $\psi(x)\to 1/2$ as $x\to 0$, (3.9) is equal to

$$\frac{1}{2}\,\bar E[P^{\lambda}_{\beta_+,\varepsilon}\{\gamma_\lambda^{-1}(\Gamma_\lambda)\}] + \Delta'(\varepsilon,\lambda)$$

which, by (3.6), is again equal to

$$(3.10)\qquad \frac{1}{2}\,\bar E[\tilde P_{\lambda\beta_+}\{\gamma_\lambda^{-1}(\Gamma_\lambda)\}] + \Delta'(\varepsilon,\lambda) + \Delta''(\varepsilon,\lambda)\ ,$$

where $\Delta'(\varepsilon,\lambda)\to 0$ as $\varepsilon\downarrow 0$ uniformly in λ and $\Delta''(\varepsilon,\lambda)\to 0$ as $\lambda\to\infty$ for each fixed $\varepsilon > 0$. Therefore, from (3.7) $\sim$ (3.10) we have

$$(3.11)\qquad I_\lambda = \frac{1}{2}\,\bar E[\tilde P_{\lambda\beta_+}\{\gamma_\lambda^{-1}(\Gamma_\lambda)\}] + o(1),\quad \lambda\to\infty\ ,$$

and a similar formula for II_λ.

To complete the proof of the theorem we need the following simple lemma.

<u>*Lemma 4*</u>. *For any* $\lambda > 0$ *and* $\Gamma\in\mathcal{B}$ *we have*

$$\tilde P_{\lambda\beta_+}\{\gamma_\lambda^{-1}(\Gamma)\} \overset{d}{=} \tilde P_{\beta_+}\{\Gamma\}\ .$$

Proof. For each W with $e^{-W}\in L^1(\mathbb{R})$, $\tilde X(t,W)$ denotes the stationary $\mathcal{L}_W$-diffusion process with initial distribution $\tilde\mu_W$. Fixing a sample path $\beta_+(\cdot)$, we put $\tilde X(t) = \tilde X(t,\beta_+(\cdot))$ and $\tilde X_\lambda(t) = \tilde X(t,\beta_+(\lambda^2\cdot))$. Since the scaling relation (3.5) implies

$$\tilde P_{\lambda\beta_+}\{\gamma_\lambda^{-1}(\Gamma)\} \overset{d}{=} \tilde P_{\beta_+(\lambda^2\cdot)}\{\gamma_\lambda^{-1}(\Gamma)\}\ ,$$

for the proof of the lemma it is enough to show

$$\{\gamma_\lambda\tilde X_\lambda(t),\ t\ge 0\} \overset{d}{=} \{\tilde X(t),\ t\ge 0\}$$

for each fixed sample path $\beta_+(\cdot)$. In what follows the notation $\hat P$

stands for the image measure of $\tilde{\mu}_{\beta_+(\cdot)}$ under the map $\hat{S}: \mathbb{R} \to \mathbb{R}$, where $\hat{S}(x) = \int_0^x e^{\beta_+(y)}dy$. Similarly $\hat{\mu}_\lambda$ stands for the image measure obtained by replacing $\beta_+(\cdot)$ by $\beta_+(\lambda^2\cdot)$. The following fact can be easily verified:

(3.12) If $B(0)$ is a random variable with distribution $\hat{\mu}_\lambda$, then $\lambda^2 B(0)$ is distributed according to $\hat{\mu}$.

As in 1.1, $\tilde{X}_\lambda(t)$ can be constructed from a Brownian motion $B(t)$ with initial distribution $\hat{\mu}_\lambda$. Write $B(t) = B(0) + \tilde{B}(t)$ and put $B_\lambda(t) = B(0) + \lambda^{-2}\tilde{B}(\lambda^4 t)$. Then $B_\lambda(t)$ is again a Brownian motion with initial distribution $\hat{\mu}_\lambda$. Therefore if, in the construction of $\tilde{X}_\lambda(t)$, we use $B_\lambda(t)$ instead of $B(t)$, we still have a diffusion process $\hat{X}_\lambda(t)$ which is equivalent in law to $\tilde{X}_\lambda(t)$. By an easy calculation we see that

$$\begin{aligned}\hat{X}_\lambda(t) &= \lambda^{-2}\hat{S}^{-1}(\lambda^2 B(0) + \tilde{B}(\hat{A}^{-1}(\lambda^4 t))) ,\\ &= \lambda^{-2}\hat{S}^{-1}(\hat{B}(\hat{A}^{-1}(\lambda^4 t))) ,\end{aligned}$$

where $\hat{B}(t) = \lambda^2 B(0) + \tilde{B}(t)$ is a Brownian motion with initial distribution $\hat{\mu}$ (by (3.12)) and $\hat{A}^{-1}(t)$ is the inverse function of

$$\hat{A}(s) = \int_0^s e^{-2\beta_+(\hat{S}^{-1}(\hat{B}(u)))}du .$$

Therefore, $\{\gamma_\lambda \hat{X}_\lambda(t)\}$ and consequently $\{\gamma_\lambda \tilde{X}_\lambda(t)\}$ is equivalent in law to $\{\tilde{X}(t)\}$.

The proof of the theorem is now completed as follows. Take an arbitrary $\Gamma \in \mathcal{B}_t$, $t > 0$ being arbitrarily fixed, and put $\Gamma_\lambda = \Gamma$ in (3.2). Then from (3.2), (3.11) and Lemma 4 we have

$$\lim_{\lambda\to\infty} E_Q[P_W\{\theta_{\exp\lambda}\omega - \lambda^2 b(W_\lambda) \in \Gamma\}] = \tilde{E}[\tilde{P}_{\beta_+}\{\Gamma\}] ,$$

and this prove the theorem.

REFERENCES

[1] Th. Brox, A one-dimensional diffusion process in a Wiener medium, Ann. Probab., 14(1986), 1206-1218.

[2] A. O. Golosov, Localization of random walks in one-dimensional random environments, Commun. Math Phys., 92(1984), 491-506.

[3] N. Ikeda and S. Watanabe, Stochastic Differential Equations and Diffusion Processes, North-Holland, 1981.

4] K. Itô and H. P. McKean, Diffusion Processes and Their Sample Paths, Springer-Verlag, 1965.

5] K. Itô, Poisson point processes attached to Markov Processes, Proc. 6th Berkeley Symp. Math. Statist. Probab. III, 225-239, Univ. California Press, Berkeley, 1972.

6] K. Kawazu, Y. Tamura and H. Tanaka, One-dimensional diffusions and random walks in random environments, to appear in Proc. 5th Japan-USSR Symp. Probab. Th.

7] H. Kesten, The limit distribution of Sinai's random walk in random environment, Physica, 138A(1986), 299-309.

8] P. Lévy, Processus stochastiques et mouvement brownien, Gauthier-Villars, Paris, 1948.

9] J. W. Pitman, One-dimensional Brownian motion and the three-dimensional Bessel process, Adv. Appl. Probab., 7(1975), 511-526.

10] J. W. Pitman and M. Yor, A decomposition of Bessel bridges, Z. Wahrscheinlichkeitstheorie Verw. Gebiete, 59(1982), 425-457.

11] Y. G. Sinai, The limiting behavior of a one-dimensional random walk in a random medium, Theory of Probab. Appl., 27(1982), 256-268.

12] H. Tanaka, Limit distributions for one-dimensional diffusion processes in self-similar random environments, to appear in Hydrodynamic Behavior and Interacting Particle Systems, the IMA Volumes in Math. and its Appl., Vol. 9, 1987, Springer-Verlag.

13] D. Williams, Path decomposition and continuity of local time for one-dimensional diffusions, I, Proc. London Math. Soc., (3)28(1974), 738-768.

Reprinted from

Stochastic Analysis, 156–172,

Lecture Notes in Math., **1322**, Springer-Verlag, 1988.

On the maximum of a diffusion process in a drifted Brownian environment

KIYOSHI KAWAZU AND HIROSHI TANAKA

1. Introduction

In this paper we investigate asymptotic behavior of the tail of the distribution of the maximum of a diffusion process in a drifted Brownian environment. This problem is a diffusion analogue of the Afanas'ev problem([1]). Our result is naturally compatible with that of Afanas'ev[1].

Let $\{W(x), x \in \mathbf{R}, P\}$ be a Brownian environment, namely, let $\{W(t), t \geq 0, P\}$ and $\{W(-t), t \geq 0, P\}$ be independent Brownian motions in one-dimension with $W(0) = 0$. We consider a diffusion process $X(t, W)$ defined formally by

$$X(t, W) = \text{Brownian motion} - \frac{1}{2}\int_0^t \{W'(X(s, W)) + c\} ds,$$

where c is a positive constant. The precise meaning of $X(t, W)$ is simply a diffusion process with generator

$$\frac{1}{2}e^{W(x)+cx}\frac{d}{dx}(e^{-W(x)-cx}\frac{d}{dx}),$$

starting at 0. Such a diffusion process can be constructed from a Brownian motion through changes of scale and time. For a fixed environment $W = (W(x), x \in \mathbf{R})$ we denote by P_W the probability law of the process $\{X(t, W)\}$ and put

$$\mathcal{P} = \int P(dW) P_W.$$

Thus $\mathcal{P}$ is the full law of $\{X(t, \cdot)\}$. We often write $X(t) = X(t, \cdot)$. Since $c > 0$, $\max_{t\geq 0} X(t)$ is finite ($\mathcal{P}$-a.s.). The problem is the following : How fast does $\mathcal{P}\{\max_{t\geq 0} X(t) > x\}$ decay as $x \to \infty$? Since

$$\mathcal{P}\{\max_{t\geq 0} X(t) > x\} = E\{A(A+B)^{-1}\}, \tag{1.1}$$

where

$$(1.2) \qquad A = \int_{-\infty}^{0} e^{W(t)+ct}\,dt, \quad B = \int_{0}^{x} e^{W(t)+ct}\,dt,$$

the problem is nothing but to find the asymptotics of $E\{A(A+B)^{-1}\}$ as $x \to \infty$. The result varies according as $c > 1$, $c = 1$, $0 < c < 1$, as will be stated in the following theorem.

THEOREM. (i)*If* $c > 1$, *then*

$$\mathcal{P}\{\max_{t\geq 0} X(t) > x\} \sim \frac{2c-2}{2c-1}\exp\{-(c-\frac{1}{2})x\}, \quad x \to \infty.$$

(ii) If $c = 1$, *then*

$$\mathcal{P}\{\max_{t\geq 0} X(t) > x\} \sim (2/\pi)^{1/2} x^{-1/2} \exp\{-x/2\}, \quad x \to \infty.$$

(iii) If $0 < c < 1$, *then*

$$\mathcal{P}\{\max_{t\geq 0} X(t) > x\} \sim const. x^{-3/2} \exp\{-c^2 x/2\}, \quad x \to \infty,$$

where

$$const. = 2^{5/2-2c}\Gamma(2c)^{-1} \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty z(a+z)^{-1} a^{2c-1} e^{-a/2} y^{2c} e^{-\lambda z} u \sinh u \, da\, dy\, dz\, du,$$

$\lambda = (1+y^2)/2 + y\cosh u$.

2. Proof of the theorem

Since A and B are independent, the right hand side of (1.1) equals $E\{Af(A)\}$ where $f(a) = E\{(a+B)^{-1}\}, a \geq 0$. Fixing $x > 0$, we consider the time reversal $\widehat{W}(t) = W(x-t) - W(x), 0 \leq t \leq x$. Since $\{\widehat{W}(t), 0 \leq t \leq x\}$ is also a Brownian motion, we have

$$(2.1) \qquad \begin{aligned} f(a) &= E\{(a + \int_0^x \exp\{\widehat{W}(t) + ct\}\,dt)^{-1}\} \\ &= E\{(a + e^{-W(x)}\int_0^x \exp\{W(x-t)+ct\}\,dt)^{-1}\} \\ &= E\{(ae^{W(x)-cx} + \int_0^x e^{W(t)-ct}\,dt)^{-1} e^{W(x)-cx}\} \\ &= e^{(1/2-c)x} E\{(ae^{W(x)-cx} + \int_0^x e^{W(t)-ct}\,dt)^{-1} e^{W(x)-x/2}\} \\ &= e^{(1/2-c)x} E\{(ae^{W(x)-(c-1)x} + \int_0^x e^{W(t)-(c-1)t}\,dt)^{-1}\} \\ &= e^{(1/2-c)x} E\{(a + \int_0^x e^{W(t)+(c-1)t}\,dt)^{-1} e^{W(x)+(c-1)x}\}. \end{aligned}$$

In deriving the fifth equality in the above we used the formula of Cameron-Martin-Maruyama-Girsanov; the last equality was derived by using $\widehat{W}(t)$ as in the case of the first equality. From the fifth equality of (2.1) we obtain the following lemma.

LEMMA 1. *For any $c > 0$ and $x > 0$*

$$\mathcal{P}\{\max_{t\geq 0} X(t) > x\} = e^{(1/2-c)x} E\{A(Ae^{W(x)-(c-1)x} + \int_0^x e^{W(t)-(c-1)t} dt)^{-1}\}, \tag{2.2}$$

where A is given by (1.2).

The following lemma due to Yor will also be used.

LEMMA 2(Yor[2]). *For any $\nu > 0$ we have*

$$\int_0^\infty \exp(W(t) - \frac{\nu t}{2}) dt \stackrel{d}{=} 2/Z_\nu, \tag{2.3}$$

where $\stackrel{d}{=}$ means equality in distribution and Z_ν is a gamma variable of index ν , that is,

$$P\{Z_\nu \in dt\} = \Gamma(\nu)^{-1} t^{\nu-1} e^{-t} dt \qquad (t > 0).$$

2.1. Proof of (i)

When $c > 1$, Lemma 1 implies

$$\lim_{x\to\infty} e^{-(1/2-c)x} \mathcal{P}\{\max_{t\geq 0} X(t) > x\} = E\{A(\int_0^\infty e^{W(t)-(c-1)t} dt)^{-1}\} \quad .$$

It is easy to see that the above expectation is finite. To obtain its exact value we use Lemma 2. We thus obtain (i).

2.2. Proof of (ii)

For $x > 0$ we put

$$\varphi(x) = E\{\log \int_0^x e^{W(t)} dt\}, \quad \psi(x) = \frac{d}{dx}\varphi(x).$$

Then it is easy to see that

$$\psi(x) = E\{(\int_0^x e^{W(t)} dt)^{-1} e^{W(x)}\} = E\{(\int_0^x e^{W(t)} dt)^{-1}\} \ ;$$

in fact, the second equality is a consequence of the last equality of (2.1) with $a = 0$ and $c = 1$. Thus $\psi(x)$ is monotone decreasing in x .

LEMMA 3. *When $c = 1$, we have*

$$E\{A(\int_0^x e^{W(t)+t} dt)^{-1}\} \sim \sqrt{2/\pi}\, x^{-1/2} e^{-x/2} \quad as \quad x \to \infty \quad . \tag{2.4}$$

Proof. Since $E\{A\} = 2$ in case $c = 1$, the left hand side of (2.4) equals $2E\{(\int_0^x e^{W(t)+t} dt)^{-}$ which also equals $2e^{-x/2} E\{(\int_0^x e^{W(t)} dt)^{-1} e^{W(x)}\}$ by virtue of (2.1) with $a = 0$ and $c = 1$. Thus we have

$$E\{A(\int_0^x e^{W(t)+t} dt)^{-1}\} = 2e^{-x/2}\psi(x). \tag{2.5}$$

On the other hand, using the scaling property $\{W(t)\} \stackrel{d}{=} \{\sqrt{x}\,W(t/x)\}$ we have

$$\varphi(x) = E\{\log \int_0^1 e^{\sqrt{x}W(t)}\,dt\} + \log x\,,$$

and hence

$$\begin{aligned} \lim_{x\to\infty} x^{-1/2}\varphi(x) &= \lim_{x\to\infty} E\{\frac{1}{\sqrt{x}}\log \int_0^1 e^{\sqrt{x}W(t)}\,dt\} \\ &= E\{\max_{0\le t\le 1} W(t)\} = \sqrt{2/\pi}\,, \end{aligned}$$

which combined with the monotonicity of $\psi(x) = \varphi'(x)$ implies

$$\psi(x) \sim (2\pi x)^{-1/2} \quad \text{as } x \to \infty\quad . \tag{2.6}$$

This together with (2.5) proves the lemma.

LEMMA 4. *For $x > 0$ we have*

$$E\{(\int_0^x e^{W(t)}\,dt)^{-2} e^{W(x)}\} \le \psi(x/2)^2 \quad . \tag{2.7}$$

Proof. The left hand side of (2.7) is dominated by

$$\begin{aligned} &E\{(\int_0^{x/2} e^{W(t)}\,dt)^{-1}(\int_{x/2}^x e^{W(t)}\,dt)^{-1} e^{W(x)}\} \\ =\ &E\{(\int_0^{x/2} e^{W(t)}\,dt)^{-1}(\int_{x/2}^x e^{W(t)-W(x/2)}\,dt)^{-1} e^{W(x)-W(x/2)}\} \\ =\ &E\{(\int_0^{x/2} e^{W(t)}\,dt)^{-1}\} E\{(\int_0^{x/2} e^{W(t)}\,dt)^{-1} e^{W(x/2)}\} \\ =\ &\psi(x/2)^2\ ; \end{aligned}$$

in deriving the second equality in the above we used the fact that $\{W(t+\frac{x}{2})) - W(\frac{x}{2}), t \ge 0\}$ is a Brownian motion independent of $\{W(t), 0 \le t \le x/2\}$.

The proof of (ii) is now given as follows. By (1.1) we have

$$\begin{aligned} 0\ &\le E\{A(\int_0^x e^{W(t)+t}\,dt)^{-1}\} - \mathcal{P}\{\max_{t\ge 0} X(t) > x\} \\ &= E\{AB^{-1} - A(A+B)^{-1}\} \\ &\le E\{2^{-1}A^{3/2}B^{-3/2}\} = 2^{-1}E\{A^{3/2}\}E\{B^{-3/2}\}. \end{aligned} \tag{2.8}$$

We prove

$$E\{A^{3/2}\} < \infty\,, \tag{2.9}$$

$$E\{B^{-3/2}\} < \text{const.}\, x^{-3/4} e^{-x/2}\,. \tag{2.10}$$

(2.9) follows immediately from Lemma 2 ; a direct proof can also be given as follows. Using Hölder's inequality we have

$$\begin{aligned} E\{A^{3/2}\} &= E\{(\int_0^\infty e^{W(t)-4t/5}e^{-t/5}\,dt)^{3/2}\} \\ &\le (5/3)^{1/2}E\{\int_0^\infty \exp\{\frac{3}{2}(W(t)-\frac{4t}{5})\}dt\} = (5/3)^{1/2}\cdot(40/3). \end{aligned}$$

(2.10) can be proved by making use of the CMMG formula, the Schwarz inequality, Lemma 4 and then (2.6) ; in fact, putting $B_0 = \int_0^x e^{W(t)}dt$ we have

$$\begin{aligned} E\{B^{-3/2}\} &= E\{B_0^{-3/2}e^{W(x)-x/2}\} \\ &\le e^{-x/2}E\{B_0^{-1}e^{W(x)}\}^{1/2}E\{B_0^{-2}e^{W(x)}\}^{1/2} \\ &\le e^{-x/2}\psi(x)^{1/2}\psi(x/2) \\ &\le \text{const.}\, e^{-x/2}x^{-1/4}\cdot x^{-1/2}. \end{aligned}$$

The assertion (ii) of our theorem follows from Lemma 3, (2.8), (2.9) and (2.10).

2.3. Proof of (iii)

The proof of (iii) relies essentially on the following Yor's formula.

Yor's formula([3: the formula(6.e)]). *For any bounded Borel functions f and g we have*

$$\begin{aligned} &E\{f(\int_0^t e^{2W(s)}\,ds)g(e^{W(t)})\} \\ &= c_t\int_0^\infty dy\int_0^\infty dz\, g(y)f(1/z)\exp\{-z(1+y^2)/2\}\psi_{yz}(t), \end{aligned}$$

where

$$\begin{aligned} c_t &= (2\pi^2 t)^{-1/2}\exp\{\pi^2/2t\}, \\ \psi_r(t) &= \int_0^\infty \exp\{-u^2/2t\}e^{-r(\cosh u)}(\sinh u)\sin(\pi u/t)\,du. \end{aligned}$$

To proceed to the proof of (iii) we put

$$f(a,z) = a(a+4z)^{-1},\quad g(y) = y^{2c},$$

$$B^{(\nu)}(t) = \int_0^t e^{2(W(s)+\nu s)}\,ds.$$

Using first the CMMG formula and then Yor's formula we have

$$\begin{aligned} &E\{a(a+\int_0^x e^{W(t)+ct}\,dt)^{-1}\} = E\{a(a+4B^{(2c)}(x/4))^{-1}\} \\ &= E\{a(a+4B^{(0)}(x/4))^{-1}\exp(2cW(x/4)-\tfrac{c^2x}{2})\} \\ &= \exp(-c^2x/2)E\{f(a,B^{(0)}(x/4))g(e^{W(x/4)})\} \\ &= \exp(-c^2x/2)c_{x/4}\int_0^\infty dy\int_0^\infty dz\, g(y)f(a,1/z)\exp\{-z(1+y^2)/2\}\psi_{yz}(x/4). \end{aligned}$$

Since Lemma 2 implies

$$P\{A \in da\} = 2^{2c}\Gamma(2c)^{-1}a^{-2c-1}e^{-2/a}da \quad (a > 0),$$

we have

$$\begin{aligned} &\mathcal{P}\{\max_{t\geq 0} X(t) > x\} \\ &= 2^{2c+1/2}\Gamma(2c)^{-1}\pi^{-1}\exp(2\pi^2/x)x^{-1/2}\exp(-c^2x/2) \\ &\times \int_0^\infty dy \int_0^\infty dz \int_0^\infty du\, y^{2c}h(z)e^{-\lambda z}\exp(-2u^2/x)(\sinh u)\sin(4\pi u/x), \end{aligned} \tag{2.11}$$

where

$$\begin{aligned} h(z) &= \int_0^\infty az(az+4)^{-1}a^{-2c-1}e^{-2/a}\,da, \\ \lambda &= (1+y^2)/2 + y\cosh u. \end{aligned}$$

LEMMA 5. *Let* $0 < c < 1$ *and put*

$$F(y,z,u) = y^{2c}h(z)e^{-\lambda z}u\sinh u.$$

Then we have

$$M = \int_0^\infty \int_0^\infty \int_0^\infty F(y,z,u)\,dy\,dz\,du < \infty,$$

Proof. By a change of variable $\cosh u = v$, we have

$$M = \int_0^\infty dy \int_0^\infty dz \int_1^\infty dv\; y^{2c}h(z)e^{-\lambda z}\log(v+\sqrt{v^2-1}),$$

where $\lambda = (1+y^2)/2 + yv$. Since

$$h(z) = 2^{-2c-1}z\int_0^\infty u^{2c-1}e^{-u}(u+\frac{z}{2})^{-1}du,$$

it is easy to see that

$$h(z) \longrightarrow 2^{-2c}\Gamma(2c) \qquad \text{as } z\to\infty \quad , \tag{2.12}$$

$$h(z) \sim_{\text{as } z\downarrow 0} \begin{cases} 2^{-2c-1}\Gamma(2c-1)z & \text{if } c > 1/2, \\ 2^{-2}z\log 1/z & \text{if } c = 1/2, \\ 2^{-4c}\int_0^\infty a^{2c-1}(a+1)^{-1}da\cdot z^{2c} & \text{if } 0 < c < 1/2. \end{cases} \tag{2.13}$$

Therefore for any $\varepsilon > 0$ and $\alpha > 0$ we have

$$\begin{aligned} M_1 &= \int_0^\infty dy \int_1^\infty dz \int_1^\infty dv\, y^{2c}h(z)e^{-\lambda z}\log(v+\sqrt{v^2-1}) \\ &\leq \text{const.} \int_0^\infty \int_1^\infty y^{2c}v^{\varepsilon}\lambda^{-1}e^{-\lambda}\,dydv \\ &\leq \text{const.} \int_0^\infty \int_1^\infty y^{2c}v^{\varepsilon}\lambda^{-\alpha}\,dydv \\ &\leq \text{const.} \int_0^\infty \int_0^\infty y^{2c-\varepsilon-1}(1+y^2)^{-\alpha+1+\varepsilon}z^{\varepsilon}(1+z)^{-\alpha}\,dy\,dz \end{aligned}$$

(by putting $v = (2y)^{-1}(1+y^2)z$ with y fixed),

which is finite if $\varepsilon > 0$ is sufficiently small and $\alpha > 0$ sufficiently large. Note that const. in the above may vary from place to place and depend on ε and α . Next we prove that

$$(2.14) \qquad M_2 = \int_0^\infty dy \int_0^1 dz \int_1^\infty dv\; y^{2c} h(z) e^{-\lambda z} \log(v + \sqrt{v^2-1}) < \infty \quad .$$

Assume $1/2 < c < 1$. Then by (2.13)

$$\begin{aligned} M_2 &\le \text{const.} \int_0^\infty dy \int_0^1 dz \int_1^\infty dv\, y^{2c} z e^{-\lambda z} v^\varepsilon \\ &\le \text{const.} \int_0^\infty \int_1^\infty \lambda^{-2} y^{2c} v^\varepsilon \; dy\, dv \qquad \left(\text{ we used } \int_0^1 z e^{-\lambda z} dz \le \lambda^{-2}\ \right) \\ &\le \text{const.} \int_0^\infty \int_0^\infty y^{2c-1-\varepsilon}(1+y^2)^{-1+\varepsilon} z^\varepsilon (1+z)^{-2}\, dy\, dz \end{aligned}$$

(by putting $v = (2y)^{-1}(1+y^2)z$ with y fixed)

which is finite for sufficiently small $\varepsilon > 0$ by virtue of $1/2 < c < 1$. When $c = 1/2$, (2.13) implies

$$M_2 \le \text{const.} \int_0^\infty dy \int_0^1 dz \int_1^\infty dv\; y z^{1-\varepsilon} e^{-\lambda z} v^\varepsilon$$

for $0 < \varepsilon < 1$. Since $\int_0^1 z^{1-\varepsilon} e^{-\lambda z}\, dz \le \text{const.}\lambda^{-2+\varepsilon}$, we have

$$\begin{aligned} M_2 &\le \text{const.} \int_0^\infty \int_1^\infty \lambda^{-2+\varepsilon} y v^\varepsilon \, dy\, dv \\ &\le \text{const.} \int_0^\infty \int_0^\infty y^{-\varepsilon}(1+y^2)^{-1+2\varepsilon} z^\varepsilon (1+z)^{-2+\varepsilon}\, dy\, dz < \infty \end{aligned}$$

provided that $\varepsilon > 0$ is small enough. Finally assume $0 < c < 1/2$. Then by (2.13)

$$\begin{aligned} M_2 &\le \text{const.} \int_0^\infty dy \int_0^1 dz \int_1^\infty dv\, y^{2c} z^{2c} e^{-\lambda z} v^\varepsilon \\ &\le \text{const.} \int_0^\infty \int_1^\infty \lambda^{-1-2c} y^{2c} v^\varepsilon \, dy\, dv \\ &\le \text{const.} \int_0^\infty \int_0^\infty y^{2c-\varepsilon-1}(1+y^2)^{-2c+\varepsilon} z^\varepsilon (1+z)^{-1-2c}\, dy\, dz < \infty \end{aligned}$$

provided that $\varepsilon > 0$ is small enough. Thus (2.14) is proved.

We can now complete the proof of (iii) as follows. From (2.11) we have

$$(2.15) \qquad \mathcal{P}\{\max_{t\ge 0} X(t) > x\} = 2^{2c+5/2}\Gamma(2c)^{-1} \exp(2\pi^2/x) x^{-3/2} \exp(-c^2 x/2) M(x)\,,$$

where

$$M(x) = \int_0^\infty \int_0^\infty \int_0^\infty F(y,z,u) \sin(4\pi u/x)/(4\pi u/x) \exp(-2u^2/x)\, dy\, dz\, du\,.$$

By Lemma 5 we have $\lim_{x\to\infty} M(x) = M$ which equals

$$2^{-4c}\int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty z(a+z)^{-1}a^{2c-1}e^{-a/2}y^{2c}e^{-\lambda z}u\sinh u\,da\,dy\,dz\,du\,.$$

Thus the assertion (iii) follows from (2.15).

Acknowledgment. We wish to thank Prof. S.Kotani and Prof. M.Yor for giving us valuable information ; Prof. M.Yor kindly sent us preprints including [2] and [3], without which the result (iii) would not have been obtained.

References

[1] V.I.Afanas'ev, On a maximum of a transient random walk in random environment, Theor.Probab.Appl.**35**(1990), 205 - 215.

[2] M.Yor, Sur certaines fonctionelles exponentielles du mouvement brownien réel, J.Appl.Probab.**29**(1992) , 202 - 208.

[3] M.Yor, On some exponential functionals of Brownian motion, to appear in Adv.Appl. Probab. (September 1992).

Kiyoshi Kawazu
Department of Mathematics
Faculty of Education
Yamaguchi University
Yosida, Yamaguchi 753
Japan

Hiroshi Tanaka
Department of Mathematics
Faculty of Science and Technology
Keio University
Hiyoshi, Yokohama 223
Japan

Reprinted from
Seminaire de Probabilites *XXVII*, *78–85*,
Lecture Notes in Math., **1557**, Springer-Verlag, 1993.

86. *Recurrence of a Diffusion Process in a Multidimensional Brownian Environment**)

By Hiroshi TANAKA

Department of Mathematics, Faculty of Science and Technology, Keio University
(Communicated by Kiyosi ITÔ, M. J. A., Nov. 12, 1993)

Introduction. Let $\boldsymbol{W}$ be the space of continuous functions on $\boldsymbol{R}^d$ vanishing at the origin. In this paper an element of $\boldsymbol{W}$ is called an environment. Given an environment W, we consider a diffusion process $\boldsymbol{X}_W = \{X(t),\ t \geq 0,\ P_W^x,\ x \in \boldsymbol{R}^d\}$ with generator

$$\frac{1}{2}(\Delta - \nabla W \cdot \nabla) = \frac{1}{2} e^W \sum_{k=1}^{d} \frac{\partial}{\partial x_k}\left(e^{-W} \frac{\partial}{\partial x_k}\right).$$

When W is bounded, the result of Nash [8] for fundamental solutions of parabolic equations guarantees the existence of a diffusion process $\boldsymbol{X}_W^0$ with generator

$$\sum_{k=1}^{d} \frac{\partial}{\partial x_k}\left(e^{-W} \frac{\partial}{\partial x_k}\right).$$

For a general W we still have a nice diffusion process $\boldsymbol{X}_W^0$ (e.g. see [4]) and hence $\boldsymbol{X}_W$ can be constructed from $\boldsymbol{X}_W^0$ through a random time change. Without any assumption on the behavior of $W(x)$ for large $|x|$ the process $\boldsymbol{X}_W$ may explode within a finite time, but such a case is excluded automatically since we are interested in the recurrence of $\boldsymbol{X}_W$. We consider the probability measure P on $\boldsymbol{W}$ with respect to which $\{W(x),\ x \in \boldsymbol{R}^d,\ P\}$ is a Lévy's Brownian motion with a d-dimensional time. The collection of diffusion processes $\boldsymbol{X} = \{\boldsymbol{X}_W\}$ in which W is allowed to vary as a random element in $(\boldsymbol{W}, P)$ is called a diffusion in a d-dimensional Brownian environment. When $d = 1$ this was considered by Brox [1] and Schumacher [9] as a diffusion model exhibiting the same asymptotic behavior as Sinai's random walk in a random environment ([10]); see also [11] for some refined results. Recently Mathieu [7] obtained some very interesting results concerning a long time asymptotic problem for $\boldsymbol{X}$ in the case $d \geq 2$. Motivated by [7] the present paper was written.

In this paper we prove that $\boldsymbol{X}_W$ is recurrent for almost all Brownian environments W in any dimension d, namely, for any nonnegative Borel function f on $\boldsymbol{R}^d$ such that $f > 0$ on a set of positive Lebesgue measure the equality

$$P_W^x\left\{\int_0^\infty f(X(t))\,dt = \infty\right\} = 1,\ x \in \boldsymbol{R}^d,$$

holds for almost all W with respect to P. In [3] Fukushima, Nakao and Takeda discussed the same problem but with the replacement of $W(x)$ by $\tilde{W}(|x|)$,

*) This research was partially supported by Grant-in-Aid for Science Research No. 04452011, the Ministry of Education, Science and Culture, Japan.

where $\tilde{W}(t)$ is a Brownian motion with a 1-dimensional time.

To obtain our result we employ Ichihara's recurrence criterion ([4]) which, in the present special case, asserts that X_W^0 (W is fixed) is recurrent if

$$\int_1^\infty \left\{\int_{S^{d-1}} e^{-W(r\theta)} d\theta\right\}^{-1} r^{-d+1} dr = \infty,$$

where $d\theta$ is the uniform distribution on S^{d-1}. We can also employ Fukushima's recurrence criterion ([2]) which, in the present special case, asserts that X_W (W is fixed) is recurrent if there exists a sequence $\{u_n\}$ such that $0 \leq u_n \leq 1$, $\lim u_n = 1$ a.e. and $\lim \mathscr{E}(u_n, u_n) = 0$, where $\mathscr{E}(u, v)$ is the Dirichlet form associated with X_W, namely,

$$\mathscr{E}(u, v) = \frac{1}{2}\int_{R^d} \nabla u \cdot \nabla u\, e^{-W} dx.$$

Since it is obvious that X_W is recurrent if and only if X_W^0 is recurrent, either criterion yields our result. But for the verification of these criteria we need some information on the asymptotic behavior of $W(x)$ for large $|x|$. A key point in obtaining this informaiton is to consider the one-parameter family $\{T_t, t \in \boldsymbol{R}\}$ of measure preserving transformations on $(\boldsymbol{W}, P)$ defined by (1. 3) and then to use its ergodicity.

§1. Brownian motion with a d-dimensional time. Let $d \geq 2$ and as before let P be the probability measure on $\boldsymbol{W}$ such that $\{W(x), x \in \boldsymbol{R}^d, P\}$ is a Brownian motion with a d-dimensional time ([6: p. 277]), that is, a Gaussian system with

$$E\{W(x)\} = 0, \ W(0) = 0, \tag{1.1}$$

$$E\{W(x)W(y)\} = \frac{1}{2}\{|x| + |y| - |x - y|\}. \tag{1.2}$$

For each $t \in \boldsymbol{R}$ and $W \in \boldsymbol{W}$ we define an element $T_t W$ of $\boldsymbol{W}$ by

$$(T_t W)(x) = e^{-t/2} W(e^t x), \quad x \in \boldsymbol{R}^d. \tag{1.3}$$

Then $\{T_t, t \in \boldsymbol{R}\}$ is a one-parameter family of measure preserving transformations on the probalility space $(\boldsymbol{W}, P)$. Using (1. 2) we can easily compute the covariance matrix of

$$e^{-t/2}W(e^t x_1), e^{-t/2}W(e^t x_2), \cdots, e^{-t/2}W(e^t x_m), W(x_1'), W(x_2'), \cdots, W(x_n')$$

for fixed $t \in \boldsymbol{R}$ and $x_1, \ldots, x_m, x_1', \ldots, x_n' \in \boldsymbol{R}^d$, and the following lemma can be proved in the same way as in Itô [5].

Lemma 1. *$\{T_t, t \in \boldsymbol{R}\}$ is mixing and hence ergodic.*

Next let $0 < a < b$, put $K = \{x \in \boldsymbol{R}^d : a \leq |x| \leq b\}$ and consider the Banach space $B = C(K)$, the space of real valued continuous functions on K, and the real Hilbert space $H = L^2(K, dx)$. The inner product in H is denoted by $\langle \cdot, \cdot \rangle$. Regarding $W_K = \{W(x), x \in K\}$ as an H-valued random variable, we denote by γ the probability distribution of W_K. Since every Borel set in the space B is also a Borel set in the space H and since W_K is regarded as a B-valued random variable, we have $\gamma(B) = 1$. γ is a Gaussian measure on H with

$$\int_H e^{\langle f, g \rangle} \gamma(dg) = E\left\{\exp \int_K f(x) W(x) dx\right\} = \exp\left\{\frac{1}{2}\langle Af, f \rangle\right\}, f \in H, \tag{1.4a}$$

(1.4b) $$Af(x) = \int_K \frac{1}{2} \{|x| + |y| - |x-y|\} f(y) dy.$$

For $\phi = Af_0$ with $f_0 \in H$ we define the ϕ-transform γ_ϕ by $\gamma_\phi(\Gamma) = \gamma(\{g : g + \phi \in \Gamma\})$. Then the following Cameron-Martin formula is easily verified by using (1. 4).

(1.5) $$\gamma_\phi(dg) = \exp\left\{\langle f_0, g\rangle - \frac{1}{2}\langle Af_0, f_0\rangle\right\}\gamma(dg).$$

Lemma 2. *Any nonempty open set in the space B has a positive γ-*measure.

Proof. We first prove that the range $R = \{Af : f \in H\}$ is dense in B. If this were not true, there exists a finite signed measure $\mu \neq 0$ on K such that

(1.6) $$\int_K Af(x)\mu(dx) = 0 \quad \text{for all} \quad f \in H.$$

Since the left hand side of (1. 6) equals $\langle f, g\rangle$ where

$$g(x) = \int_K \frac{1}{2} \{|x| + |y| - |x-y|\} \mu(dy) \in H,$$

(1.6) implies $g = 0$. Therefore, regarding μ as a signed measure in $\boldsymbol{R}^d$ we have

(1.7) $$\int_{\boldsymbol{R}^d}\int_{\boldsymbol{R}^d} \frac{1}{2} \{|x| + |y| - |x-y|\} \mu(dy)\mu(dy) = 0.$$

In the same way as in the proof of Théorème 58 of [6: p. 276] we can prove that the left hand side of (1. 7) equals

$$\text{const.} \int_{\boldsymbol{R}^d} |\xi|^{-d-1} |\hat{\mu}(\xi) - \hat{\mu}(0)|^2 d\xi,$$

where $\hat{\mu}(\xi)$ is the Fourier transform of μ. Therefore μ must be concentrated on $\{0\}$. But this is impossible because $0 \notin K$ and hence R must be dense in B. Next we notice that the whole space B, which has γ-measure 1, can be expresses as a union of a countable number of open balls of the form $B_\varepsilon(\phi) = \{\psi \in B : \|\psi - \phi\|_\infty < \varepsilon\}$, $\phi \in R$, $\varepsilon > 0$ being arbitrary but fixed. On the other hand by the Cameron-Martin formula (1.5) $\gamma(B_\varepsilon(\phi)) = \gamma_{-\phi}(B_\varepsilon(0)) > 0$ if and only if $\gamma(B_\varepsilon(0)) > 0$ provided that $\phi \in R$. Therefore we must have $\gamma(B_\varepsilon(\phi)) > 0$ for all $\phi \in R$. This implies the assertion of the lemma.

§2. Recurrence of X_W. Since our result in the 1-dimensional case is easily obtained from a general theory of 1-dimensional diffusion processes, we assume $d \geq 2$.

Theorem 1. *X_W is recurrent for almost all Brownian environments W.*

Proof. It is enough to prove that X_W^0 is recurrent for almost all Brownian environments W and, according to Ichihara's criterion ([4: Theorem A]) it is also enough to prove that

(2.1) $$\int_1^\infty \left\{\int_{S^{d-1}} e^{-W(r\theta)} d\theta\right\}^{-1} r^{-d+1} dr = \infty, \ P\text{-}a.s.$$

If we put $M(t) = \min\{(T_t W)(\theta) : \theta \in S^{d-1}\}$, then

(2.2) the left hand side of (2. 1)

$$= \int_0^\infty e^{(2-d)t} \left\{ \int_{S^{d-1}} \exp(-e^{t/2}(T_t W)(\theta))d\theta \right\}^{-1} dt$$

$$\geq \int_0^\infty e^{(2-d)t} \exp\{e^{t/2} M(t)\} dt \geq \int_0^\infty \mathbf{1}_{(a,\infty)}(M(t))dt,$$

provided that $a > 0$ is chosen so that $(2-d)t + ae^{t/2} \geq 0$ holds for all $t \geq 0$. Next take $K = \{x \in \boldsymbol{R}^d : 1 \leq |x| \leq 2\}$ and consider B, H and γ as in the preceeding section. Since $\Gamma = \{\phi \in B : \min(\phi(x) : |x| = 1) > a\}$ is an open set in B, we have $\gamma(\Gamma) > 0$ by Lemma 2. The ergodicity of $\{T_t, t \in \boldsymbol{R}\}$ now implies

$$\lim_{T\to\infty} T^{-1} \int_0^T \mathbf{1}_{(a,\infty)}(M(t))dt = E\{\mathbf{1}_{(a,\infty)}(M(0))\} = \gamma(\Gamma) > 0, \ P\text{-}a,s.,$$

and hence $\int_0^\infty \mathbf{1}_{(a,\infty)}(M(t))dt = \infty$, P-$a.s.$, which combined with (2.2) proves (2. 1).

Remark 1. X_W is null-recurrent (P-$a.s.$) in the sense that $m_W(dx) = e^{-W}dx$ is an invariant measure for X_W with $m_W(\boldsymbol{R}^d) = \infty$.

Remark 2. Fukushima's criterion can also be used for proving Theorem 1; in fact, by virtue of Lemmas 1, 2 it is still easy to prove the existence of a sequence of radial functions u_n in $C_0^\infty(\boldsymbol{R}^d)$ such that $0 \leq u_n \leq 1$, $\lim u_n = 1$ a.e. and $\mathscr{E}(u_n, u_n) = 0$. This argument also proves the recurrence of $X_{|W|}$ for almost all Brownian environments W.

Remark 3. $X_{-|W|}$ is recurrent for $d = 1$ and transient for $d \geq 2$ for almost all Brownian environments W. The proof in the case $d \geq 2$ is as follows. According to Theorem B of [4] the transience of $X^0_{-|W|}$ (and consequently of $X_{-|W|}$) follows if one proves that, for almost all Brownian environments W,

$$\int_0^\infty e^{-|W(r\theta)|} r^{-d+1} dr < \infty \tag{2.3}$$

for θ belonging to some subset (which may depend on W) of S^{d-1} with a positive uniform measure. But this can be proved by showing that the expectation (with respect to P) of the left hand side of (2. 3) is finite for each fixed θ.

References

[1] Th. Brox: A one-dimensional diffusion process in a Wiener medium. Ann. Probab., **14**, 1206–1218 (1986).

[2] M. Fukushima: On recurrence criteria in the Dirichlet space theory. Local Times to Global Geometry. Control and Physics (ed. Elworthy). Research Notes in Math., Series 150, Longman (1987).

[3] M. Fukushima, S. Nakao and M. Takeda: On Dirichlet forms with random data - recurrence and homogenization. Stochastic Processes - Mathematics and Physics II (eds. S. Albeverio, Ph. Blanchard and L. Streit). Lect. Notes in Math., vol. 1250, Springer-Verlag (1987).

[4] K. Ichihara: Some global properties of symmetric diffusion processes. Publ. RIMS, Kyoto Univ., **14**, 441–486 (1978).

[5] K. Itô: On the ergodicity of a certain stationary process. Proc. Imp. Acad., **20**, 54–55 (1944).

[6] P. Lévy: Processus Stochastiques et Mouvement Brownien. Gauthier-Villars, Paris (1948).

[7] P. Mathieu: Zero white noise limit through Dirichlet forms. Application to diffusions in a random medium (1992) (preprint).

[8] J. Nash: Continuity of solutions of parabolic and elliptic equations. Amer. J. Math., **80**, 931–953 (1958).

[9] S. Schumacher: Diffusions with random coefficients. Contemporary Math. Particle Systems, Random Media and Large Deviations (ed. Durrett). vol. 41, pp. 351–356 (1985).

[10] Y. G. Sinai: The limiting behavior of a one-dimensional random walk in a random medium. Theor. Probab. Appl., **27**, 256–268 (1982).

[11] H. Tanaka: Limit theorem for one-dimensional diffusion process in Brownian environment. Stochastic Analysis (Proc. Japanese-French Seminar, Paris 1987). Lect. Notes in Math., vol. 1322, Springer-Verlag, pp. 156–172.

Localization of a Diffusion Process in a One-Dimensional Brownian Environment

HIROSHI TANAKA
Keio University

Dedicated to H. P. McKean.

1. Introduction

Let $\mathbb{W}$ be the space of continuous functions $W : \mathbb{R} \to \mathbb{R}$ with $W(0) = 0$ and Q be the Wiener measure on $\mathbb{W}$, namely, let $\{W(x), x \geqq 0, Q\}$ and $\{W(-x), x \geqq 0, Q\}$ be independent one-dimensional Brownian motions starting at 0. An element of $\mathbb{W}$ is called an environment. Given an environment W we consider a diffusion process $X(t, W)$ starting at 0 and with generator

$$L_W = \frac{1}{2} e^{W(x)} \frac{d}{dx} \left(e^{-W(x)} \frac{d}{dx} \right) . \tag{1.1}$$

$X(t, W)$ can be constructed from a one-dimensional Brownian motion $B(t)$ by a scale change and a time change; see [3]. We denote by (Ω, P) the probability space on which $B(t)$ is defined and put $\mathcal{P} = P \otimes Q$. Thus $B(t)$ and $W(x)$ are independent. $X(t, W)$ is then called a diffusion process in a Brownian environment. Sinai presented this model as a diffusion analogue of his random walk in a random environment; see [10], page 268. Brox (see [1]) and Schumacher (see [9]) proved that $X(t, \cdot)$ exhibits an asymptotic behavior as $t \to \infty$ which is similar to that of Sinai's random walk of [10]. We consider another probability measure $\widetilde{Q}$ on $\mathbb{W}$ such that $\{W(x), x \geqq 0, \widetilde{Q}\}$ and $\{W(-x), x \geqq 0, \widetilde{Q}\}$ are independent Bessel processes of index 3 starting at 0. Then $e^{-W} \in L^1(\mathbb{R}), \widetilde{Q}$-a.s. For each W with $e^{-W} \in L^1(\mathbb{R})$ let μ_W be the probability measure in $\mathbb{R}$ of the form

$$\mu_W(dx) = \text{const. } e^{-W(x)} dx . \tag{1.2}$$

For each $\lambda > 0$ we define $b_\lambda(W)$, a function of the environment W alone, by (3.1). Then it was proved in [11] that the distribution of $X(e^\lambda, \cdot) - b_\lambda(\cdot)$ converges to μ as $\lambda \to \infty$, where $\mu = \int \mu_W \widetilde{Q}(dW)$. A similar localization theorem had already been obtained by Golosov (see [2]) for reflecting random walks on $\mathbb{Z}^+$. In the present paper we restate the above localization theorem for $X(t, \cdot)$ in a modified form and prove it by a method different from that of [11].

THEOREM 1.1. *For any t_j, $1 \leqq j \leqq k$, with $0 < t_1 < \cdots < t_k$ the joint distribution of $X(e^\lambda t_j, \cdot) - b_\lambda(\cdot)$, $1 \leqq j \leqq k$, with respect to $\mathcal{P}$ converges to the mixture $\int \mu_W^k \widetilde{Q}(dW)$ as $\lambda \to \infty$ where μ_W^k is the k-fold product distribution of μ_W.*

Communications on Pure and Applied Mathematics, Vol. XLVII, 755–766 (1994)

CCC 0010-3640/94/050755-12

Ogura (see [7], Example 7.2) gave another proof to Brox's Theorem 1.4 (see [1]) from which the present method was suggested. The notion of valleys of environments still plays an essential role as before. The main difference between the present proof and that of [11] is as follows: In [11] we used Brox's estimates for exit times from valleys and a coupling technique, but now, instead of these, we use a theorem of Ogura (see [7]) concerning the convergence of a sequence of one-dimensional (generalized) diffusion processes.

2. Convergence Theorem for Diffusion Processes in One-Dimension

It is well known that the generator of a diffusion process in $\mathbb{R}$ can be expressed as a diffusion operator $d/\{m(dx)\}\, d/\{dS(x)\}$. The measure $m(dx)$, called the speed measure, is finite on compact subsets and positive on open subsets ($\neq \emptyset$) of $\mathbb{R}$, and the function $S(x)$, called the canonical scale, is continuous and strictly increasing. Ogura (see [7]) discussed a more general case, namely, the case of "bi-generalized diffusion processes," but for our present purpose it is enough to consider a sequence of diffusion processes. Thus, suppose we are given a sequence of diffusion operators L_n with speed measure $m_n(dx)$ and canonical scale $S_n(x)$, $n = 1, 2, \dots$, and denote by $X_n^x(t)$ the diffusion process with generator L_n starting at x. We assume that the following conditions (2.1), (2.2), and (2.3) are satisfied:

(2.1) For each n $S_n(0) = 0$ and $S_n(x)$ tends to ∞ or $-\infty$ accordingly as $x \to \infty$ or $x \to -\infty$; for each x the canonical scale $S_n(x)$ tends to 0 as $n \to \infty$.

(2.2) For any $f \in C_0(\mathbb{R})$, the space of continuous functions with compact supports,

$$\lim_{n\to\infty} \int f\, dm_n = \int f\, dm\,,$$

where m is a nontrivial finite measure (namely, m_n converges vaguely to m as $n \to \infty$).

(2.3) The measure $\widetilde{m}_n = m_n \circ S_n^{-1}$ converges vaguely to $c\delta_0$ as $n \to \infty$, where S_n^{-1} is the inverse function of S_n, $c = m(\mathbb{R}) > 0$ and δ_0 denotes the δ-measure at 0.

The author learned the following theorem through the kindness of Y. Ogura.

OGURA'S THEOREM. (SEE [7]) *Let ε be an arbitrary constant such that $0 < \varepsilon < 1$, put*

$$T_{k,\varepsilon} = \{(t_1, \dots, t_k) \in \mathbb{R}^k : \varepsilon \leqq t_1 < t_k \leqq 1/\varepsilon\,,\quad t_j - t_{j-1} \geqq \varepsilon\,(1 \leqq \forall j \leqq k)\} \tag{2.4}$$

and consider a sequence $\{x_n\}$ satisfying

$$|S_n(x_n)| \leqq \frac{1}{\varepsilon} \quad \textit{for } n = 1, 2, \dots\,. \tag{2.5}$$

Then for any $f_j \in C_0(\mathbb{R})$, $1 \leqq j \leqq k$,

$$E\left\{\prod_{j=1}^{k} f_j\left(X_n^{x_n}(t_j)\right)\right\} \to \prod_{j=1}^{k}\int c^{-1} f_j\, dm$$

as $n \to \infty$ *uniformly in* $\{x_n\}$ *satisfying condition* (2.5) *and in* $(t_1, \cdots, t_k) \in T_{k,\varepsilon}$.

Proof: Since the case of $\lim S_n(x) = 0$ is excluded in Theorem 5.1 of [7], strictly speaking the above theorem is not a straightforward consequence of Theorem 5.1 of [7]. The proof can be done, however, in the same way. Denote by $q_n(t,x,y)$ the transition density with respect to $\widetilde{m}_n$ of the diffusion process with speed measure $\widetilde{m}_n$ and canonical scale x. It is defined through formula (5.6) of [7] by taking $\widetilde{m} = \widetilde{m}_n$ and $\widetilde{Q} = \mathbb{R}$. It is known that $q_n(t,x,y)$ is continuous in $(t,x,y) \in (0,\infty) \times \mathbb{R} \times \mathbb{R}$ and symmetric in x and y. Also we define $q(t,x,y)$ through formula (5.6) of [7] by taking $\widetilde{m} = c\delta_0$ and $\widetilde{Q} = \mathbb{R}$. By an easy computation we see that $q(t,x,y) = 1/c$ for any $(t,x,y) \in (0,\infty)\times\mathbb{R}\times\mathbb{R}$. Probabilistically (and formally) $q(t,x,y)$ is the transition density with respect to $c\delta_0$ of the "generalized diffusion process" which jumps to 0 immediately at $t = 0$ and remains there for any $t > 0$. Now Proposition 5.1 of [7] implies that $q_n(t,x,y)$ converges to $q(t,x,y)$ uniformly on each compact subset of $(0,\infty)\times\mathbb{R}\times\mathbb{R}$ as $n \to \infty$. Therefore, if $p_n(t,x,y)$ denotes the transition density with respect to m_n of $X_n(t)$, then for any $f \in C_0(\mathbb{R})$

$$\int p_n(t,x_n,y) f(y) m_n(dy)$$
$$= \int q_n(t, S_n(x_n), S_n(y)) f(y) m_n(dy) \to \int c^{-1} f\, dm$$

as $n \to \infty$ uniformly in $t \geqq \varepsilon$ and in $\{x_n\}$ satisfying condition (2.5), $\varepsilon > 0$ being an arbitrary constant. This implies the theorem.

3. Valleys

3.1. We give the definition of a valley of an environment W following [1]. A part $\{W(x),\, x_1 \leqq x \leqq x_2\}$ of W is called a *valley* of W if:

(i) $x_1 < x_2$;

(ii) there exists $x_0 \in (x_1, x_2)$ such that

$$W(x_1) > W(x) > W(x_0) \quad \text{for any } x \in (x_1, x_0)\,,$$
$$W(x_2) > W(x) > W(x_0) \quad \text{for any } x \in (x_0, x_2)\,;$$

(iii) for the same x_0 as above

$$H_- = \sup\{W(y) - W(x) : x_1 \leqq x < y \leqq x_0\} < W(x_2) - W(x_0)\,,$$
$$H_+ = \sup\{W(x) - W(y) : x_0 \leqq x < y \leqq x_2\} < W(x_1) - W(x_0)\,.$$

For simplicity we write (x_1, x_0, x_2) instead of $\{W(x), x_1 \leqq x \leqq x_2\}$ and call x_0 the *bottom* of the valley. $A = H_+ \vee H_-$ is called the *inner directed ascent* (abbreviated to *i.d.a.*) and $D = (W(x_1) - W(x_0)) \wedge (W(x_2) - W(x_0))$ is the *depth* of the valley. A valley (x_1, x_0, x_2) is said to contain 0 if $x_1 < 0 < x_2$. For any fixed $r > 0$, with Q-measure 1, there exists a valley of W with *i.d.a.* $< r <$ *depth* and containing 0 (see [1]); moreover, there are many such valleys for fixed W but it can be easily seen that their bottoms are the same. We denote by $b = b(W)$ the unique (common) bottom of such valleys for $r = 1$.

To give another description of a valley due to Kesten (see [6]) let $\lambda > 0, W \in \mathbb{W}$ and put

$$W^{\#}(x) = \begin{cases} W(x) - \min\limits_{[0,x]} W & \text{for } x \geqq 0\,, \\ W(x) - \min\limits_{[x,0]} W & \text{for } x < 0\,, \end{cases}$$

$$d_\lambda^+ = \min\{x > 0 : W^{\#}(x) = \lambda\}\,, \qquad V_\lambda^+ = \min_{[0,d_\lambda^+]} W\,,$$

$$d_\lambda^- = \max\{x < 0 : W^{\#}(x) = \lambda\}\,, \qquad V_\lambda^- = \min_{[d_\lambda^-,0]} W\,.$$

We define b_λ^+ and b_λ^- by $W(b_\lambda^\pm) = V_\lambda^\pm$ (such $b_\lambda^\pm$ are uniquely determined with Q-measure 1 for each fixed $\lambda > 0$). Let

$$M_\lambda^+ = \max_{[0,b_\lambda^+]} W\,, \qquad J_\lambda^+ = M_\lambda^+ \vee (V_\lambda^+ + \lambda)\,,$$

$$M_\lambda^- = \max_{[b_\lambda^-,0]} W\,, \qquad J_\lambda^- = M_\lambda^- \vee (V_\lambda^- + \lambda)\,,$$

and define $b_\lambda = b_\lambda(W)$ by

$$b_\lambda(W) = \begin{cases} b_\lambda^+ & \text{if } J_\lambda^+ < J_\lambda^-\,, \\ b_\lambda^- & \text{if } J_\lambda^+ > J_\lambda^-\,. \end{cases} \tag{3.1}$$

Moreover, we define $a_\lambda = a_\lambda(W)$ and $c_\lambda = c_\lambda(W)$ by

(3.2) $a_\lambda(W)$ = the infimum of the set of a's ($a < b_\lambda$) such that

$$\begin{cases} W(a) > W(x) > W(b_\lambda) & \text{for any } x \in (a, b_\lambda)\,, \\ \sup\{W(y) - W(x) : a \leqq x < y \leqq b_\lambda\} < \lambda\,, \end{cases}$$

(3.3) $c_\lambda(W)$ = the supremum of the set of c's ($c > b_\lambda$) such that

$$\begin{cases} W(c) > W(x) > W(b_\lambda) & \text{for any } x \in (b_\lambda, c)\,, \\ \sup\{W(x) - W(y) : b_\lambda \leqq x < y \leqq c\} < \lambda\,. \end{cases}$$

Then we see that, with Q-measure $1, (a_\lambda, b_\lambda, c_\lambda)$ is the maximum valley of W containing 0 and satisfying $A_\lambda < \lambda < D_\lambda$, where A_λ and D_λ are, respectively, the i.d.a. and the depth of the valley. Thus b_1 coincides with b of the preceding paragraph. From now on we suppress the suffix 1 in $a_1, b_1, c_1, D_1, d_1^+, d_1^-$, etc.

3.2. For $\lambda > 0$ and $W \in \mathbb{W}$ we define $W_\lambda \in \mathbb{W}$ by $W_\lambda(x) = \lambda^{-1}W(\lambda^2 x)$, $x \in \mathbb{R}$. Then for each fixed $\lambda > 0$ $\{W_\lambda, Q\}$ is equivalent in law to $\{W, Q\}$ and $a_\lambda(W) = \lambda^2 a(W_\lambda), b_\lambda(W) = \lambda^2 b(W_\lambda), c_\lambda(W) = \lambda^2 c(W_\lambda), D_\lambda(W) = \lambda D(W_\lambda)$. As stated in the Introduction we denote by $X(t, W)$ the diffusion process with generator L_W and starting at 0. Then the following scaling relation holds (see [1], Lemma 1.3): For any fixed $\lambda > 0$ and $W \in \mathbb{W}$

$$\{X(t, \lambda W_\lambda), t \geqq 0, P\} \stackrel{d}{=} \{\lambda^{-2}X(\lambda^4 t, W), t \geqq 0, P\} \tag{3.4}$$

where $\stackrel{d}{=}$ means the equality in distribution (W is fixed).

3.3. The following is a result of Brox, proved in [1], Proposition 4.1 (see also [4], Theorem I-A-1): *If (x_1, x_0, x_2) is a valley of W with i.d.a. $< r <$ depth and containing 0, then $X(e^{\lambda r}, \lambda W)$ converges to x_0 in probability as $\lambda \to \infty$ (W is fixed).*

3.4. On a suitable probability space we consider a process $\{w^+(x), x \in \mathbb{R}\}$ such that $\{w^+(x), x \geqq 0\}$ and $\{w^+(-x), x \geqq 0\}$ are independent reflecting Brownian motions on $[0, \infty)$ starting at 0. Let $\{\ell^+(x), x \in \mathbb{R}\}$ be the local time at 0 of $w^+(x)$, namely,

$$\ell^+(x) = \lim_{\varepsilon \downarrow 0} \left(\frac{1}{2\varepsilon}\right) \int_I \mathbf{1}_{[0,\varepsilon]}(w^+(y))\, dy \,,$$

where $I = [0, x]$ or $[x, 0]$ accordingly as $x \geqq 0$ or $x < 0$. Now we put $\beta^+(x) = w^+(x) + \ell^+(x)$. Then by Pitman's theorem (see [8]) $\{\beta^+(x), x \geqq 0\}$ and $\{\beta^+(-x), x \geqq 0\}$ are independent Bessel processes of index 3 starting at 0. For $\lambda > 0$ we put

(3.5a) $\sigma_\lambda^+ =$ the smallest zero of $w^+(x)$ in $(z, 0]$ where z is the maximum of

$$x < 0 \quad \text{with } w^+(x) = \lambda \,.$$

(3.5b) $\tau_\lambda^+ = \min\{x > 0 : \beta^+(x) = \lambda\}$.

We next introduce another process $\{w^-(x), x \in \mathbb{R}\}$ which is equivalent in law to $\{w^+(x), x \in \mathbb{R}\}$. We assume that $w^+(x)$ and $w^-(x)$ are defined on a common probability space $(\bar{\Omega}, \bar{P})$ and that they are independent. Denote by $\ell^-(x)$ the local time at 0 of $w^-(x)$, put $\beta^-(x) = w^-(x) + \ell^-(x)$ and define σ_λ^- and τ_λ^- in a way similar to equation (3.5). We put

$$\begin{aligned} \bar{b}_\lambda^+ &= -\sigma_\lambda^+ , & \bar{b}_\lambda^- &= \sigma_\lambda^- , \\ \bar{d}_\lambda^+ &= \tau_\lambda^+ - \sigma_\lambda^+ , & \bar{d}_\lambda^- &= -(\tau_\lambda^- - \sigma_\lambda^-) , \\ \bar{V}_\lambda^+ &= -\beta^+(\sigma_\lambda^+) , & \bar{V}_\lambda^- &= -\beta^-(\sigma_\lambda^-) , \\ \bar{M}_\lambda^+ &= \max_{[\sigma_\lambda^+,0]} \beta^+ - \beta^+(\sigma_\lambda^+) , & \bar{M}_\lambda^- &= \max_{[\sigma_\lambda^-,0]} \beta^- - \beta^-(\sigma_\lambda^-) , \\ \bar{J}_\lambda^+ &= \bar{M}_\lambda^+ \vee (\bar{V}_\lambda^+ + \lambda) , & \bar{J}_\lambda^- &= \bar{M}_\lambda^- \vee (\bar{V}_\lambda^- + \lambda) . \end{aligned}$$

The following lemma for $\lambda = 1$ was proved in [11]. The proof for a general $\lambda > 0$ can be done in the same way.

LEMMA 3.1. (SEE [11]) *Let $\lambda > 0$ and put*

$$\begin{aligned} W_\lambda^+ &= \left\{ W(x + b_\lambda^+) - W(b_\lambda^+) , \quad -b_\lambda^+ \leqq x \leqq d_\lambda^+ - b_\lambda^+ \right\} , \\ W_\lambda^- &= \{ W(-x + b_\lambda^-) - W(b_\lambda^-) , \quad b_\lambda^- \leqq x \leqq -(d_\lambda^- - b_\lambda^-) \} . \end{aligned}$$

Then, under the law Q $(W_\lambda^+, J_\lambda^+)$ and $(W_\lambda^-, J_\lambda^-)$ are independent and have the same distribution which is equal to the joint distribution of $\{\beta^+(x), \sigma_\lambda^+ \leqq x \leqq \tau_\lambda^+\}$ and $\bar{J}_\lambda^+$.

LEMMA 3.2. *Let $\widetilde{W}_\lambda(x) = \lambda\{W(\lambda^{-2}x + b) - W(b)\}, x \in \mathbb{R}$. Then $\{\widetilde{W}_\lambda, Q\}$ converges in law to $\{W, \widetilde{Q}\}$ as $\lambda \to \infty$.*

Proof: It is enough to prove that, for any positive constant K and for any bounded continuous function F on $\mathbb{W}$ depending only on $\{W(x), |x| \leqq K\}$,

$$\lim_{\lambda \to \infty} \int F\left(\widetilde{W}_\lambda\right) Q(dW) = \int F\, d\widetilde{Q} . \tag{3.6}$$

From now on we suppress the suffix 1 in J_1^+, J_1^+, W_1^+, etc. We put $\Gamma_\lambda = \{J^+ < J^-\} \cap \{\lambda^{-2}K < b^+ \wedge (d^+ - b^+)\}$ and notice that $Q(\Gamma_\lambda)$ increases to $Q\{J^+ < J^-\} = 1/2$ as $\lambda \uparrow \infty$. Moreover, if $W \in \Gamma_\lambda$ then the value of $F(\widetilde{W}_\lambda)$ is determined by W^+, namely, we can write

$$F\left(\widetilde{W}_\lambda\right) = G_\lambda(W^+) \quad \text{on } \Gamma_\lambda$$

with a suitable function G_λ. Thus, using the notation $E_Q\{F;\Gamma\} = \int_\Gamma F dQ$ and applying Lemma 3.1, we can write

$$
\begin{aligned}
&E_Q\left\{F\left(\widetilde{W}_\lambda\right);\ J^+ < J^-\right\} \\
&= E_Q\left\{F\left(\widetilde{W}_\lambda\right);\ \Gamma_\lambda\right\} + o(1) \\
&= E_Q\{G_\lambda(W^+);\ \Gamma_\lambda\} + o(1) \\
&= \bar{E}\left\{G_\lambda\left(\beta^+(x), \sigma^+ \leqq x \leqq \tau^+\right);\ \bar{\Gamma}_\lambda\right\} + o(1) \\
&= \bar{E}\left\{F\left(\lambda\beta^+\left(\lambda^{-2}\cdot\right)\right);\ \bar{\Gamma}_\lambda\right\} + o(1) \\
&= \bar{E}\left\{F\left(\lambda\beta^+(\lambda^{-2}\cdot)\right);\ \bar{J}^+ < \bar{J}^-\right\} + o(1),
\end{aligned}
\tag{3.7}
$$

where $\bar{\Gamma}_\lambda = \{\bar{J}^+ < \bar{J}^-\} \cap \{\lambda^{-2}K < \bar{b}^+ \wedge (\bar{d}^+ - \bar{b}^+)\}$, and $o(1)$ represents a term which tends to 0 as $\lambda \to \infty$. By an argument similar to that used for arriving at expression (3.9) of [11" we can prove that the last term in equation (3.7) is asymptotically equal to

$$
\bar{E}\left\{F\left(\lambda\beta^+\left(\lambda^{-2}\cdot\right)\right)\right\}/2 \tag{3.8}
$$

as $\lambda \to \infty$. Since $\{\lambda\beta^+(\lambda^{-2}\cdot), \bar{P}\}$ is equivalent in law to $\{\beta^+(\cdot), \bar{P}\}$, (3.8) is equal to

$$
\bar{E}\{F(\beta^+(\cdot))\}/2 = \int F d\widetilde{Q}/2\,, \tag{3.9}
$$

and hence the first term of equation (3.7) tends to (3.9) as $\lambda \to \infty$. Since $E_Q\{F(\widetilde{W}_\lambda); J^+ > J^-\}$ can be treated similarly, we obtain equality (3.6).

4. Proof of Theorem 1.1

The idea of the proof is to apply Ogura's theorem to study the limiting behavior of $X(e^\lambda t, \cdot) - b_\lambda(\cdot)$ as $\lambda \to \infty$.

4.1. We consider the diffusion processes

$$
X_\lambda^W(t) = X\left(e^\lambda t, W\right) - b_\lambda(W)\,, \qquad t \geqq 0\,, \tag{4.1}
$$

$$
Y_\lambda^W(t) = \lambda^2\left\{X\left(\lambda^{-4}e^\lambda t, \lambda W\right) - b(W)\right\}\,, \quad t \geqq 0\,. \tag{4.2}
$$

The scaling relation (3.4) implies that, for each fixed $\lambda > 0$ and W,

$$
\left\{X_\lambda^W(t), t \geqq 0\right\} \stackrel{\mathrm{d}}{=} \left\{Y_\lambda^{W_\lambda}(t), t \geqq 0\right\}.
$$

Since the distributions of W and W_λ under Q are the same,

$$\left\{Y_\lambda^{W_\lambda}(t), t \geqq 0\right\} \stackrel{\mathscr{D}}{=} \left\{Y_\lambda^W(t), t \geqq 0\right\},$$

where $\stackrel{\mathscr{D}}{=}$ means the equality in distribution (W is random). Thus the processes (4.1) and (4.2) are equivalent in law under $\mathscr{P}$. For a fixed W the generator of the diffusion process $Y_\lambda^W(t)$ is expressed as the diffusion operator L_λ^W with the following canonical scale and speed measure:

$$S_\lambda^W(x) = 2e^{-\lambda} \int_0^x \exp\left\{\lambda\left(W\left(\lambda^{-2}y + b\right) - W(b)\right)\right\} dy . \tag{4.3}$$

$$m_\lambda^W(dx) = \exp\left\{-\lambda\left(W\left(\lambda^{-2}x + b\right) - W(b)\right)\right\} dx . \tag{4.4}$$

LEMMA 4.1. (i) *For any x the canonical scale $S_\lambda^W(x)$ tends to 0 as $\lambda \to \infty$ with Q-measure* 1.

(ii) *We regard m_λ^W and $\widetilde{m}_\lambda^W = m_\lambda^W \circ (S_\lambda^W)^{-1}$ as random variable taking values in the space of non-negative Radon measures in $\mathbb{R}$ equipped with the topology of vague convergence. Then the joint distribution (under Q) of m_λ^W and $\widetilde{m}_\lambda^W$ converges to the joint distribution (under $\widetilde{Q}$) of $e^{-W(x)}dx$ and const.δ_0(const. $= \int e^{-W}dx$) as $\lambda \to \infty$.*

Proof: The assertion (i) is obvious. To prove (ii) we first put

$$x_1 = \max\{x < 0 : W(x+b) - W(b) = 1\} ,$$
$$x_2 = \min\{x > 0 : W(x+b) - W(b) = 1\} ,$$

and observe that, with Q-measure 1, for any $\varepsilon > 0$ there exist $x' \in (x_1 - \varepsilon, x_1)$ and $x'' \in (x_2, x_2 + \varepsilon)$ such that $W(x' + b) - W(b) > 1$ and $W(x'' + b) - W(b) > 1$. Thus, if x varies with λ in such a way that $x > \lambda^2(x_2 + \varepsilon)$ or $x < \lambda^2(x_1 - \varepsilon)$, then $|S_\lambda^W(x)| \to \infty$ as $\lambda \to \infty$. On the other hand, if x varies with λ in such a way that $\lambda^2(x_1 + \varepsilon) < x < \lambda^2(x_2 - \varepsilon)$, then $S_\lambda^W(x) \to 0$ as $\lambda \to \infty$, where $\varepsilon > 0$ can be chosen arbitrarily under the condition $\varepsilon < (-x_1) \wedge x_2$. Therefore, for each fixed $y \neq 0$ $\lambda^{-2}(S_\lambda^W)^{-1}(y)$ tends as $\lambda \to \infty$ to x_1 or x_2 accordingly as $y < 0$ or $y > 0$. Thus, putting $z_i = (S_\lambda^W)^{-1}(y_i)$, $i = 1, 2$, for any given y_1 and y_2 with $y_1 < 0 < y_2$, we see that $\lambda^{-2}z_1 \to x_1$ and $\lambda^{-2}z_2 \to x_2$ as $\lambda \to \infty$. Now we write

$$\begin{aligned}
\widetilde{m}_\lambda^W([y_1, y_2]) &= m_\lambda^W([z_1, z_2]) \\
&= \int_{z_1}^{z_2} \exp\left\{-\lambda\left(W\left(\lambda^{-2}x + b\right) - W(b)\right)\right\} dx \\
&= I_1 + I_2 + I_3 ,
\end{aligned}$$

where

$$I_1 = \int_{-1/\varepsilon}^{1/\varepsilon}, \quad I_2 = \int_{-\lambda^2\varepsilon}^{-1/\varepsilon} + \int_{1/\varepsilon}^{\lambda^2\varepsilon}, \quad I_3 = \int_{z_1}^{-\lambda^2\varepsilon} + \int_{\lambda^2\varepsilon}^{z_2},$$

and prove the following:

(4.5) $\{m_\lambda^W, Q\}$ converges in law to $\{e^{-W(x)}dx, \widetilde{Q}\}$ as $\lambda \to \infty$. In particular, for each $\varepsilon > 0\,\{I_1, Q\}$ converges in law to $\{\int_{|x|<1/\varepsilon} \exp\{-W\}dx, \widetilde{Q}\}$ as $\lambda \to \infty$.

(4.6) For any $\gamma > 0$ $\lim_{\varepsilon\downarrow 0} \overline{\lim}_{\lambda\to\infty} Q\{I_i > \gamma\} = 0$, $i = 2, 3$.

(4.5) is an immediate consequence of Lemma 3.2. For the proof of (4.6) with $i = 3$ it is enough to note that

$$\begin{aligned} I_3 &= \lambda^2 \left(\int_{\lambda^{-2}z_1}^{-\varepsilon} + \int_{\varepsilon}^{\lambda^{-2}z_2} \right) \exp\{-\lambda(W(x+b) - W(b))\}\, dx \\ &\sim \lambda^2 \left(\int_{x_1}^{-\varepsilon} + \int_{\varepsilon}^{x_2} \right) \exp\{-\lambda(W(x+b) - W(b))\}\, dx \\ &\to 0\,, \quad \lambda \to \infty\,, \end{aligned}$$

provided that W is in the event $\{(-x_1) \wedge x_2 > \varepsilon\}$ whose Q-measure tends to 1 as $\varepsilon \downarrow 0$. As for (4.6) with $i = 2$, by replacing W by W_λ we see that I_2 is equivalent in law to

$$\left(\int_{-\lambda^2\varepsilon}^{-1/\varepsilon} + \int_{1/\varepsilon}^{\lambda^2\varepsilon} \right) \exp\{-(W(x+b_\lambda) - W(b_\lambda))\}\, dx\,,$$

which is dominated by

$$\begin{aligned} &\left(\int_{-b_\lambda^+}^{-1/\varepsilon} + \int_{1/\varepsilon}^{d_\lambda^+ - b_\lambda^+} \right) \exp\left\{-(W(x+b_\lambda^+) - W(b_\lambda^+))\right\}\, dx \\ &+ \left(\int_{b_\lambda^-}^{-1/\varepsilon} + \int_{1/\varepsilon}^{-(d_\lambda^- - b_\lambda^-)} \right) \exp\{-(W(-x+b_\lambda^-) - W(b_\lambda^-))\}\, dx\,, \end{aligned} \tag{4.7}$$

provided that W is in the event

$$\begin{aligned} &\{d_\lambda^+ - b_\lambda^+ > \lambda^2\varepsilon > 1/\varepsilon\,,\ -b_\lambda^+ < -\lambda^2\varepsilon\,, \\ &-(d_\lambda^- - b_\lambda^-) > \lambda^2\varepsilon\,,\ b_\lambda^- < -\lambda^2\varepsilon\}\,. \end{aligned} \tag{4.8}$$

But by virtue of Lemma 3.1, (4.7) is equivalent in law to

$$\begin{aligned} &\left(\int_{\sigma_\lambda^+}^{-1/\varepsilon} + \int_{1/\varepsilon}^{\tau_\lambda^+} \right) \exp\{-\beta^+(x)\}\, dx \\ &+ \left(\int_{\sigma_\lambda^-}^{-1/\varepsilon} + \int_{1/\varepsilon}^{\tau_\lambda^-} \right) \exp\{-\beta^-(x)\}\, dx\,, \end{aligned}$$

which is dominated by

$$\int_{|x|>1/\varepsilon} \exp\{-\beta^+(x)\}\, dx + \int_{|x|>1/\varepsilon} \exp\{-\beta^-(x)\}\, dx\,. \tag{4.9}$$

(4.9) obviously tends to 0 as $\varepsilon \downarrow 0$, $\bar{P}$ -a.s. Since Q-measure of the event (4.8) equals Q-measure of the event $\{d^+ - b^+ > \varepsilon, b^+ > \varepsilon, b^- - d^- > \varepsilon, b^- < -\varepsilon\}$, which clearly tends to 1 as $\varepsilon \downarrow 0$, we obtain (4.6) for $i = 2$. The assertion (ii) of Lemma 4.1 now follows from (4.5) and (4.6).

4.2. We now proceed to the final part of the proof of Theorem 1.1. For any f_j and t_j, $1 \leqq j \leqq k$, with $f_j \in C_0(\mathbb{R})$ and $0 < t_1 < \cdots < t_k$ we are going to prove that

$$\int E\left\{\prod_{j=1}^{k} f_j\left(X_\lambda^W(t_j)\right)\right\} Q(dW) \to \int U(W)\widetilde{Q}(dW) \tag{4.10}$$

as $\lambda \to \infty$, where

$$U(W) = \prod_{j=1}^{k} \int f_j \, d\mu_W .$$

By the equivalence in law of the processes (4.1) and (4.2), (4.10) is equivalent to

$$\int E\left\{\prod_{j=1}^{k} f_j\left(Y_\lambda^W(t_j)\right)\right\} Q(dW) \to \int U(W)\widetilde{Q}(dW) . \tag{4.11}$$

Let $\{\lambda_n, n = 1,2,\dots\}$ be any positive sequence such that $\lambda_n \to \infty$ as $n \to \infty$. Lemma 4.1 enables us to use Skorohod's realization theorem of almost sure convergence. Thus, on a suitable probability space which we still denote by $(\bar{\Omega}, \bar{P})$, we can choose a sequence $\{\bar{W}_n\}$ and $\bar{W}$ in such a way that the following hold:

(4.12) $\{\bar{W}_n, \bar{P}\}$ is equivalent in law to $\{W, Q\}$.

(4.13) $\{\bar{W}, \bar{P}\}$ is equivalent in law to $\{W, \widetilde{Q}\}$.

(4.14) m_n converges vaguely to $e^{-\bar{W}}dx$ while $\widetilde{m}_n$ converges vaguely to const.δ_0 (const. $= \int e^{-\bar{W}}dx$) as $n \to \infty, \bar{P}$-a.s., where m_n and $\widetilde{m}_n$ are defined below.

The measure m_n is defined by equation (4.4) but with the replacement of W and λ by $\bar{W}_n$ and λ_n. $\widetilde{m}_n = m_n \circ S_n^{-1}$ where S_n is defined by (4.3) with the replacement of W and λ by $\bar{W}_n$ and λ_n. Now, fixing the value of $\bar{W}_n$, we denote by $Y_n^x(t, \bar{W}_n)$ the diffusion process with canonical scale S_n and speed measure m_n and starting at x (the probability space on which $Y_n^x(t, \bar{W}_n)$ is defined is still chosen as (Ω, P)). Then (4.14) and (i) of Lemma 4.1 imply that, with probability 1, the conditions (2.1), (2.2), and (2.3) are satisfied. (Strictly speaking, as for condition (2.1) we must take a suitable subsequence of $\{\lambda_n\}$, but this subsequence is still denoted by $\{\lambda_n\}$.) Thus Ogura's theorem is applicable to $Y_n^x(t, \bar{W}_n)$: Putting

$$g_n\left(t_1, \dots, t_k, x, \bar{W}_n\right) = E\left\{\prod_{j=1}^{k} f_j\left(Y_n^x\left(t_j, \bar{W}_n\right)\right)\right\},$$

we have

$$\lim_{n\to\infty} g_n(t_1,\ldots,t_k,x_n,\bar{W}_n) = U(\bar{W}) \ , \quad \bar{P}\text{-a.s.},$$

provided that $\{S_n(x_n)\}$ is bounded. In proving (4.11), however, a difficulty arises since the initial value of the process (4.2) (with W and λ replaced by $\bar{W}_n$ and λ_n) grows so fast with λ_n that the condition (2.5) is not satisfied. To overcome this difficulty we use Brox's result. Let $A = A(W)$ be the i.d.a. of the valley (a,b,c) of W introduced in Section 3.1. Then $A < 1, Q$-a.s. Let r be a constant such that $0 < r < 1$ and suppose $A(W) < r$. Then there exists a valley (a', b, c') of W with i.d.a. $< r <$ depth, $a \leqq a', c' \leqq c$ and containing 0. Therefore, by Brox's result stated in Section 3.3, $X(e^{\lambda r}, \lambda W)$ tends to b as $\lambda \to \infty$ in probability (W is fixed), and hence for each $\varepsilon > 0$

$$P\Big\{ \big| Y_\lambda^W(t) \big| < \lambda^2 \varepsilon \Big\} \to 1 \ , \qquad \lambda \to \infty \ , \tag{4.15}$$

where t is determined by $\lambda^{-4} e^{\lambda} t = e^{\lambda r}$, namely, $t = \lambda^4 e^{-\lambda(1-r)}$. If we denote by $\nu_n(W, dx)$, the probability distribution of $Y_\lambda^W(t)$ with $\lambda = \lambda_n, t$ as above, and W fixed, then (4.15) implies

$$\nu_n(W, (-\lambda_n^2\varepsilon, \lambda_n^2\varepsilon)) \to 1 \ , \qquad n \to \infty \ , \tag{4.16}$$

if $A(W) < r$. Next we introduce

$$\xi(W) = \max\left\{ x < 0 : W(x+b) - W(b) = \frac{1}{2} \right\},$$

$$\eta(W) = \min\left\{ x > 0 : W(x+b) - W(b) = \frac{1}{2} \right\},$$

and put $\bar{\xi}_n = \xi(\bar{W}_n), \bar{\eta}_n = \eta(\bar{W}_n)$. If $\bar{\xi}_n \wedge \bar{\eta}_n > \varepsilon$ then it is easy to see that for $0 < \varepsilon < 1/8$,

$$\big| S_n(\pm\lambda_n^2\varepsilon) \big| < 2\lambda_n^2 \int_0^\varepsilon \exp\{-\lambda_n^2/2\}\, dx < 1 \ . \tag{4.17}$$

Now, for any $\varepsilon > 0$ small but fixed we choose $r = r(\varepsilon) \in (0,1)$ so that $Q\{A < r\} > 1 - \varepsilon$ and then determine t by $\lambda^{-4} e^{\lambda} t = e^{\lambda r}$ as before. Putting $\Gamma_{n,\varepsilon} = \{A(\bar{W}_n) < r(\varepsilon), \varepsilon < \bar{\xi}_n \wedge \bar{\eta}_n\}$, we have

$$\begin{aligned} &\text{the left-hand side of (4.11)} \\ &= \int d\bar{P} \int \nu_n(\bar{W}_n, dx)\, g_n(t_1 - t, \ldots, t_k - t, x, \bar{W}_n) \\ &= \int_{\Gamma_{n,\varepsilon}} d\bar{P} \int_{(-\lambda_n^2\varepsilon, \lambda_n^2\varepsilon)} \nu_n(\bar{W}_n, dx) \\ &\quad \times g_n(t_1 - t, \ldots, t_k - t, x, \bar{W}_n) + R_{n,\varepsilon} \ . \end{aligned} \tag{4.18}$$

Since $\bar{P}\{\Gamma_{n,\varepsilon}\} = Q\{A < r(\varepsilon), \varepsilon < \xi \wedge \eta\} \to 1$ as $\varepsilon \downarrow 0$, from (4.16) we have

$$\lim_{\varepsilon \downarrow 0} \overline{\lim_{n\to\infty}} R_{n,\varepsilon} = 0 . \tag{4.19}$$

Moreover, Ogura's theorem implies that

$$\sup_{x:|S_n(x)|\leqq 1} \left| g_n(t_1 - t, \dots, t_k - t, x, \bar{W}_n) - U(\bar{W}) \right| \to 0 , \qquad \bar{P}\text{-a.s.,} \tag{4.20}$$

as $n \to \infty$. Therefore, it follows from (4.17)–(4.20) that the left-hand side of (4.11) tends to $\int U(\bar{W})d\bar{P} = \int U(W)\tilde{Q}(dW)$ as $\lambda \to \infty$ via $\{\lambda_n\}$. Since $\{\lambda_n\}$ is an arbitrary sequence with $\lambda_n \to \infty$, this implies (4.11). The proof of our theorem is finished.

Remark. In [5] a localization theorem was obtained for diffusion processes in a considerably wider class of random environments which are asymptotically self-similar. It will be possible to discuss this case by the present method but most of the essential arguments of [5] will not be simplified as much.

Bibliography

[1] Brox, T., *A one-dimensional diffusion process in Wiener medium*, Ann. Probab. 14, 1986, pp. 1206–1218.

[2] Golosov, A. O., *Localization of random walks in one-dimensional random environments*, Comm. Math. Phys. 92, 1984, pp. 491–506.

[3] Itô, K., and McKean, H. P., *Diffusion Processes and Their Sample Paths*, Springer-Verlag, New York, 1965.

[4] Kawazu, K., Tamura, Y., and Tanaka, H., *One-dimensional diffusions and random walks in random environments*, pp. 170–184 in: *Probab. Th. Math. Statist. 5th Japan-USSR Symposium Proceedings 1986*, S. Watanabe and Yu. V. Prokhorov, eds., Lecture Notes in Mathematics No. 1299, Springer-Verlag, Berlin–New York, 1988.

[5] Kawazu, K., Tamura, Y., and Tanaka, H., *Localization of diffusion processes in one-dimensional random environment*, J. Math. Soc. Japan 44, 1992, pp. 515–550.

[6] Kesten, H., *The limit distribution of Sinai's random walk in random environment*, Physica 138A, 1986, pp. 299–309.

[7] Ogura, Y., *One-dimensional bi-generalized diffusion processes*, J. Math. Soc. Japan 41, 1989, pp. 213–242.

[8] Pitman, J. W., *One-dimensional Brownian motion and the three-dimensional Bessel process*, Adv. in Appl. Probab. 7, 1975, pp. 511–526.

[9] Schumacher, S., *Diffusions with random coefficients*, pp. 351–356 in: *Particle Systems, Random Media and Large Deviations*, R. Durrett, ed., Contemporary Math. 41, 1985.

[10] Sinai, Y. G., *The limiting behavior of a one-dimensional random walk in a random medium*, Theory Probab. Appl. 27, 1982, pp. 256–268.

[11] Tanaka, H., *Limit theorem for one-dimensional diffusion process in Brownian environment*, pp. 156–172 in: *Stochastic Analysis, Proc. Japanese-French Seminar, Paris, 1987*, Lecture Notes in Mathematics No. 1322, Springer-Verlag, Berlin–New York, 1988.

Received June 1992.

Diffusion Processes in Random Environments

HIROSHI TANAKA

Department of Mathematics
Faculty of Science and Technology
Keio University
Yokohama 223, Japan

1 Introduction

Problems concerning limiting behavior of random processes in random environments have been discussed mostly in the framework of random walks (e.g., see [1], [4], [6], [15], [17], [23], [24]). Most of the problems, naturally, can also be treated in the framework of diffusion processes. We give here a survey of some recent results concerning diffusion processes in random environments, mainly of one dimension, with emphasis on the following two examples of problems. The use of methods and results in theory of diffusion processes makes our argument transparent.

(I) Localization by random centering (depending only on the environment) of a diffusion in a one-dimensional Brownian environment.
(II) Limit theorems for a diffusion in a one-dimensional Brownian environment with drift.

We also give a brief survey concerning

(III) results for a diffusion in a multidimensional Brownian environment.

2 A diffusion in a one-dimensional Brownian environment (with drift)

Let P be the Wiener measure on $\mathbf{W} = C(\mathbf{R}) \cap \{W : W(0) = 0\}$. The processes $\{W(t), t \geq 0, P\}$ and $\{W(-t), t \geq 0, P\}$ are thus independent Brownian motions. Let $\Omega = C[0, \infty)$ and denote by $\omega(t)$ the value of a function $\omega(\in \Omega)$ at time t. For fixed W and a given constant κ we consider a probability measure P_W on Ω such that $\{\omega(t), t \geq 0, P_W\}$ is a diffusion process with generator

$$\mathcal{L}_W = \frac{1}{2} e^{W(x,\kappa)} \frac{d}{dx} \left(e^{-W(x,\kappa)} \frac{d}{dx} \right)$$

and starting at 0, where $W(x, \kappa) = W(x) - \frac{1}{2}\kappa x$. It is well known that a version of $\mathbf{X}_W = \{\omega(t), t \geq 0, P_W\}$ can be obtained from a Brownian motion through a scale change and a time change. We can regard W and $\{\omega(t), t \geq 0\}$ as defined on the probability space $(\mathbf{W} \times \Omega, \mathcal{P})$ where $\mathcal{P}(dW d\omega) = P(dW) P_W(d\omega)$. The process $\mathbf{X} = \{\omega(t), t \geq, \mathcal{P}\}$ is then called a *diffusion in a Brownian environment* (*with drift*

Proceedings of the International Congress
of Mathematicians, Zürich, Switzerland 1994

if $\kappa \neq 0$). Symbolically one may write $d\omega(t) = dB(t) - \frac{1}{2}W'(\omega(t), \kappa)dt$ where $B(t)$ is a Brownian motion independent of $W(\cdot)$; however, this stochastic differential equation has no rigorous meaning.

When $\kappa = 0$, $\mathbf{X}$ is a diffusion model of Sinai's random walk in a random environment [23]. In this case Schumacher [22] and Brox [3] showed that $\mathbf{X}$ exhibits the same asymptotic behavior as Sinai's random walk, namely, that the limit distribution of $(\log t)^{-2}\omega(t)$ as $t \to \infty$ exists. Kesten [14] obtained the explicit form of the limit distribution. Golosov [7] also obtained a similar explicit form for a reflecting random walk model. Some generalizations of these results were done in [10] and [25]. The problem (I) stated in the introduction is to elaborate the result of [22] and [3] by taking account of a random centering that depends only on the environment W. This will be discussed in the next section. It is to be noted that a similar localization result was already obtained by Golosov [6] for reflecting random walks on $\mathbf{Z}^+$.

The problem (II) is concerned with the case $\kappa \neq 0$ and may be regarded as a diffusion analogue of what was discussed by Kesten-Kozlov-Spitzer [15], Solomon [24], and Afanas'ev [1]. Here we are mainly interested in limit theorems concerning the first passage time $T_x = \inf\{t > 0 : \omega(t) = x\}$ as $x \to \infty$. As will be seen in Section **4**, the result varies with κ and naturally is compatible with those of [15] and [1].

3 Localization by random centering in the case $\kappa = 0$

The argument of [3] relies on the notion of a valley introduced in [23]; in order to state only the result, however, it is adequate to start simply with the definition of the "bottom"(denoted by b_λ) of a suitable valley around the origin. Given a Brownian environment $W = \{W(x), x \in \mathbf{R}\}$, let us define $b_\lambda = b_\lambda(W)$ following [14] for each $\lambda > 0$. Setting

$$W^{\#}(x) = W(x) - \min_{[x\wedge 0, x\vee 0]} W,$$

$$d_\lambda^+ = \min\{x > 0 : W^{\#}(x) = \lambda\}, \quad V_\lambda^+ = \min_{[0,d_\lambda^+]} W,$$

$$d_\lambda^- = \max\{x < 0 : W^{\#}(x) = \lambda\}, \quad V_\lambda^- = \min_{[d_\lambda^-,0]} W,$$

we first determine b_λ^+ and b_λ^- by $W(b_\lambda^+) = V_\lambda^+$ and $W(b_\lambda^-) = V_\lambda^-$, respectively (such $b_\lambda^\pm$ are uniquely determined with P-measure 1 for each fixed $\lambda > 0$), and then define $b_\lambda = b_\lambda(W)$ by

$$b_\lambda(W) = \begin{cases} b_\lambda^+ & \text{if } M_\lambda^+ \vee (V_\lambda^+ + \lambda) < M_\lambda^- \vee (V_\lambda^- + \lambda), \\ b_\lambda^- & \text{if } M_\lambda^+ \vee (V_\lambda^+ + \lambda) > M_\lambda^- \vee (V_\lambda^- + \lambda), \end{cases}$$

where $M_\lambda^+ = \max\{W(x) : 0 \leq x \leq b_\lambda^+\}$ and $M_\lambda^- = \max\{W(x) : b_\lambda^- \leq x \leq 0\}$. When $\lambda = 1$ we write $b = b(W)$ suppressing the suffix 1. We also define $W_\lambda (\in \mathbf{W})$ for each $\lambda > 0$ and $W \in \mathbf{W}$ by $W_\lambda(x) = \lambda^{-1}W(\lambda^2 x), x \in \mathbf{R}$. Then $\{W_\lambda, P\}$ is equivalent in law to $\{W, P\}$ and hence the distribution of $b(W_\lambda)$ is independent of $\lambda > 0$.

Let $\mathbf{X} = \{\omega(t), t \geq 0, \mathcal{P}\}$ be a diffusion in a Brownian environment ($\kappa = 0$) starting at 0. According to Schumacher [22] and Brox [3]

$$\lambda^{-2}\omega(e^{\lambda}) - b(W_{\lambda}) \to 0 \tag{3.1}$$

in probability with respect to $\mathcal{P}$ *as* $\lambda \to \infty$.

Localization by random centering arises from the following question: Under what scaling does the left-hand side of (3.1) admit a nondegenerate limit distribution? The answer is simply that $\omega(e^{\lambda}) - \lambda^2 b(W_{\lambda})$ does. To state the result more precisely we need to introduce another probability measure Q on $\mathbf{W}$, defined in such a way that $\{W(x), x \geq 0, Q\}$ and $\{W(-x), x \geq 0, Q\}$ are independent Bessel processes of index 3 starting at 0. Let μ_W be the probability measure in $\mathbf{R}$ of the form $const. \exp\{-W(x)\}dx$; it is well defined for almost all W with respect to Q because $\exp(-W) \in L^1(\mathbf{R})$, Q-a.s. For an integer $k \geq 1$ we set $\mu_W^k = \mu_W \otimes \cdots \otimes \mu_W$ (the k-fold product) and $\mu^k = \int \mu_W^k Q(dW)$.

THEOREM 1 ([26], [28]). *For any* $t_1, \ldots, t_k$ *with* $0 < t_1 < \cdots < t_k$ *the joint distribution of* $\omega(e^{\lambda}t_j) - b_{\lambda}(W), 1 \leq j \leq k$, *with respect to* $\mathcal{P}$ *converges to* μ^k *as* $\lambda \to \infty$.

This theorem was proved in [26] for $k = 1$. The case $k \geq 1$ was proved in [28] by making use of Ogura's theorem stated below. Suppose we are given a sequence of diffusion operators

$$L_n = \frac{d}{m_n(dx)} \frac{d}{dS_n(x)}, \quad n \geq 1,$$

and denote by $X_n^x(t)$ the diffusion process with generator L_n starting at x. We assume that the following conditions (i), (ii), and (iii) are satisfied.

(i) $S_n(0) = 0$ and $S_n(x)$ tends to ∞ or $-\infty$ accordingly as $x \to \infty$ or $x \to -\infty$; for each x, $S_n(x) \to 0$ as $n \to \infty$.

(ii) The measure m_n converges vaguely to some nontrivial finite measure m as $n \to \infty$.

(iii) The measure $\widetilde{m_n} = m_n \circ S_n^{-1}$ converges vaguely to $c\delta_0$ as $n \to \infty$, where S_n^{-1} is the inverse function of S_n, $c = m(\mathbf{R}) > 0$, and δ_0 is the δ-measure at 0.

OGURA'S THEOREM ([21]; see also [26]). *For any* $\varepsilon \in (0,1)$ *and an integer* $k \geq 1$ *we set*

$$T_{k,\varepsilon} = \{(t_1, \ldots, t_k) \in \mathbf{R}^k : \varepsilon \leq t_1 < t_k \leq 1/\varepsilon, t_j - t_{j-1} \geq \varepsilon \ (1 \leq \forall j \leq k)\} \tag{3.2}$$

and consider a sequence $\{x_n\}$ *satisfying*

$$|S_n(x_n)| \leq 1/\varepsilon, \quad n \geq 1. \tag{3.3}$$

Then for any continuous functions f_j *in* $\mathbf{R}$ *with compact supports*, $1 \leq j \leq k$,

$$E\left\{\prod_{j=1}^{k} f_j(X_n^{x_n}(t_j))\right\} \to \prod_{j=1}^{k} \int f_j \, dm_0$$

as $n \to \infty$ *uniformly in* $\{x_n\}$ *satisfying the condition* (3.3) *and in* $(t_1, \ldots, t_k) \in T_{k,\varepsilon}$, *where* m_0 *is the probability measure* $c^{-1}m$.

It is known (see [3], Lemma 1.3) that, for fixed W, the process $\{\omega(\lambda^4 t), t \geq 0, P_W\}$ is equivalent in law to $\{\lambda^2\omega(t), t \geq 0, P_{\lambda W_\lambda}\}$. This combined with the fact that $b_\lambda(W) = \lambda^2 b(W_\lambda)$ implies that the process $\{\omega(e^\lambda t) - b_\lambda(W), t \geq 0, P_W\}$ is equivalent in law to $\{\lambda^2(\omega(\lambda^{-4}e^\lambda t) - b(W_\lambda)), t \geq 0, P_{\lambda W_\lambda}\}$; in addition, W_λ and W are identical in law. Therefore, for the proof of Theorem 1 it is enough to show

$$\int E_{\lambda W}\left\{\prod_{j=1}^{k} f_j(\lambda^2(\omega(\lambda^{-4}e^\lambda t_j) - b(W)))\right\} P(dW)$$
$$\longrightarrow \int \left\{\prod_{j=1}^{k}\int f_j\, d\mu_W\right\} Q(dW), \quad \lambda \to \infty. \tag{3.4}$$

For fixed W the generator of the diffusion process $\{\lambda^2(\omega(\lambda^{-4}e^\lambda t) - b(W)), t \geq 0, P_{\lambda W}\}$ is given by $\{d/m_\lambda^W(dx)\}\{d/dS_\lambda^W(x)\}$, where

$$S_\lambda^W(x) = 2e^{-\lambda}\int_0^x \exp\{\lambda(W(\lambda^{-2}y + b) - W(b))\}\, dy,$$
$$m_\lambda^W(dx) = \exp\{-\lambda(W(\lambda^{-2}x + b) - W(b))\}dx.$$

LEMMA 1 ([28]). (i) *$S_\lambda^W(x)$ tends to* 0 *as $\lambda \to \infty$ with P-measure* 1.
(ii) *If we regard m_λ^W and $\widetilde{m}_\lambda^W = m_\lambda^W \circ (S_\lambda^W)^{-1}$ as random variables taking values in the space of Radon measures in* **R** *equipped with the topology of vague convergence, then the joint distribution (under P) of m_λ^W and $\widetilde{m}_\lambda^W$ converges to the joint distribution (under Q) of $\exp\{-W(x)\}dx$ and $c_W\delta_0$ as $\lambda \to \infty$ where $c_W = \int \exp\{-W(x)\}\, dx$.*

Making use of Lemma 1 and Ogura's theorem we can prove (3.4) and hence Theorem 1. For details see [28].

A similar localization problem was discussed in [11] when $\{W(x)\}$ is a step process arising from a random walk that is assumed to converge in law, under a suitable scaling, to a strictly stable process.

4 Limit theorems in the case $\kappa \neq 0$

Let $\mathbf{X} = \{\omega(t), t \geq 0, \mathcal{P}\}$ denote the diffusion in a Brownian environment with drift ($\kappa \neq 0$), and set $T_x = \inf\{t > 0 : \omega(t) = x\}, \overline{\omega}(t) = \max\{\omega(s) : 0 \leq s \leq t\}$ and $\underline{\omega}(t) = \inf\{\omega(s) : s \geq t\}$.

THEOREM 2 ([13]). (i) *If $\kappa > 1$, then*

$$\lim_{x\to\infty} T_x/x = \frac{4}{\kappa - 1}, \quad \mathcal{P} - a.s.,$$
$$\lim_{t\to\infty} \omega(t)/t = \frac{\kappa - 1}{4}, \quad \mathcal{P} - a.s.$$

(ii) *If $\kappa = 1$, then $(x\log x)^{-1}T_x$ converges to* 4 *in probability* (w.r.t. $\mathcal{P}$) *as $x \to \infty$ and each of*

$$t^{-1}(\log t)\overline{\omega}(t), \quad t^{-1}(\log t)\omega(t) \quad \text{and} \quad t^{-1}(\log t)\underline{\omega}(t)$$

converges to 1/4 *in probability* (w.r.t. $\mathcal{P}$) *as $t \to \infty$.*

(iii) *If* $0 < \kappa < 1$, *then*

$$\lim_{x\to\infty} \mathcal{P}\{x^{-1/\kappa}T_x \le t\} = F_\kappa(t), \quad t > 0,$$
$$\lim_{t\to\infty} \mathcal{P}\{t^{-\kappa}\overline{\omega}(t) \le x\} = \lim_{t\to\infty} \mathcal{P}\{t^{-\kappa}\omega(t) \le x\}$$
$$= \lim_{t\to\infty} \mathcal{P}\{t^{-\kappa}\underline{\omega}(t) \le x\} = 1 - F_\kappa(x^{-1/\kappa}), \quad x > 0,$$

where F_κ *is the distribution function of a one-sided stable distribution with Laplace transform* $\exp(-c\lambda^\kappa)$.

REMARK. *The constant* c *in the Laplace transform* $\exp(-c\lambda^\kappa)$ *is given by*

$$\mathrm{c} = \left\{2^{1-\kappa}\Gamma(\kappa)\int_0^\infty \phi(x)^{-2}dx\right\}^{-1},$$

where $\phi(x)$ *is the solution of* $\frac{d}{dM(x)} \cdot \frac{d\phi}{dx} = 2\phi, \phi(0) = 1, \phi'(0) = 0; M(x)$ *is given by* $M(x) = 2\gamma(\rho^{-1}(x))$, *where* $\gamma(x) = \int_0^x z^{-\kappa}e^{-4z}dz$ *and* $\rho^{-1}(x)$ *is the inverse function of* $\rho(y) = \int_0^y z^{\kappa-1}e^{4z}\,dz$.

Theorem 2 is a diffusion analogue of (a part of) the results for random walks due to Kesten et al. [15]. We do not give a detailed proof here but we remark that our method of the proof, in particular, of (ii) and (iii) is different from that of [14] and is based on the following lemma due to Kotani.

KOTANI'S LEMMA (1988, unpublished; see [13]). *Let* $\lambda > 0$. *Then for* $t \ge 0$

$$E_W\left\{e^{-\lambda T_t}\right\} = \exp\left\{-\int_0^t U_\lambda(s)\,ds\right\}, \quad \mathcal{P}-a.s.,$$

where $U_\lambda(t)$ *is the unique stationary positive solution of*

$$dU_\lambda(t) = U_\lambda(t)dW(t) + \left\{2\lambda + \frac{1-\kappa}{2}U_\lambda(t) - U_\lambda(t)^2\right\}dt.$$

By virtue of Kotani's lemma, for the proof of (iii) it is enough to show that, with $\lambda = \xi x^{-1/\kappa}$,

$$\lim_{x\to\infty} \mathcal{E}\left\{\exp(-\xi T_x/x^{1/\kappa})\right\}$$
$$= \lim_{\lambda\downarrow 0} E\left\{\exp\left(-\int_0^{\lambda^{-\kappa}\xi^\kappa} U_\lambda(s)\,ds\right)\right\} = \exp(-c\xi^\kappa),$$

and the key point in proving the last equality is the use of Kasahara's continuity theorem [9] concerning Krein's correspondence (e.g. see [16]). A full proof is given in [13].

The following theorem is a diffusion analogue of the result of Afanas'ev [1].

THEOREM 3 ([12]). (i) *If* $-2 < \kappa < 0$, *then*

$$\mathcal{P}\{T_x < \infty\} \sim const.x^{-3/2}\exp(-\kappa^2 x/8), \quad x \to \infty,$$

where

$$const. = 2^{\frac{5}{2}+\kappa}\Gamma(-\kappa)^{-1}\int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty z(a+z)^{-1}a^{-\kappa-1}e^{-a/2}y^{-\kappa}e^{-\lambda z}u\sinh u\, da\, dy\, dz\, du,$$

$(\lambda = 2^{-1}(1+y^2) + y\cosh u)$.

(ii) *If* $\kappa = -2$, *then*

$$\mathcal{P}\{T_x < \infty\} \sim (2/\pi)^{1/2}x^{-1/2}\exp(-x/2), \quad x \to \infty.$$

(iii) *If* $\kappa < -2$, *then*

$$\mathcal{P}\{T_x < \infty\} \sim \frac{-\kappa-2}{-\kappa-1}\cdot\exp\{(\kappa+1)x/2\}, \quad x \to \infty.$$

The proof of (i) relies on an explicit representation of the distribution of a certain Brownian functional due to Yor ([29], see the formula (6.e)).

5 A diffusion in a multidimensional Brownian environment

One generalization of the model discussed in Section **3** to a multidimensional case is to take a Lévy's Brownian motion with a multidimensional time as an environment. Let $\{W(x), x \in \mathbf{R}^d, P\}$ be a Lévy's Brownian motion with a d-dimensional time that is supposed to be an environment. For a frozen Brownian environment W let $\mathbf{X}_W = \{\omega(t), t \geq 0, P_W\}$ be a diffusion process with generator $2^{-1}(\Delta - \nabla W \cdot \nabla)$ starting at 0. Existence of such a diffusion is guaranteed by the result of Nash ([20]). As in a one-dimensional case we call $\mathbf{X} = \{\omega(t), t \geq 0, \mathcal{P}\}$ a diffusion in a d-dimensional Brownian environment, where $\mathcal{P}(dW d\omega) = P(dW)P_W(d\omega)$. A similar diffusion model appeared in a heuristic argument of [18]. Durrett [4] obtained rigorous results on recurrence and localization for random walks on $\mathbf{Z}^d$ described by a certain random potential having asymptotic self-similarity and stationary increments. The diffusion $\mathbf{X}$ may be regarded as the continuous time analogue of what was discussed in Example 2 ($\beta = 1$) of [4]. Recently Mathieu [19] considered the diffusion $\mathbf{X}$ itself and discussed its long time asymptotic behavior.

THEOREM 4 ([27]; see also [4] for random walks). $\mathbf{X}_W$ *is recurrent for almost all Brownian environments* W *for any dimension* d.

This theorem can easily be proved by making use of Ichihara's recurrence test ([8], see Theorem A) concerning symmetric diffusions. We can also use Fukushima's recurrence criterion [5] in terms of the associated Dirichlet form. $\mathbf{X}_{|W|}$ is also recurrent, P-a.s.; however, $\mathbf{X}_{-|W|}$ is transient, P-a.s., for any $d \geq 2$ as can be proved by using Ichihara's transience test ([8], see Theorem B). From the argument of [27] it is also easy to see that Theorem 4 remains valid when $\{W(x)\}$ is replaced by any continuous random field $\{V(x)\}$ in $\mathbf{R}^d$ satisfying the following conditions (i), (ii), and (iii).

(i) Self-similarity: there exists $\alpha > 0$ such that the law of $\{\lambda^{-1}V(\lambda^\alpha x)\}$ equals that of $\{V(x)\}$, denoted by P, for each $\lambda > 0$.
(ii) $\{T_t, t \in \mathbf{R}\}$ is ergodic, where T_t is a P-preserving transformation from $C(\mathbf{R}^d)$ onto itself defined by $(T_tV)(x) = e^{-t/\alpha}V(e^t x), x \in \mathbf{R}^d$.
(iii) $\min\{V(x) : |x| = 1\} > 0$ with positive probability.

The argument of [19] entails the following theorem.

THEOREM 5 ([19]; see also [4] for random walks). *Localization takes place for* $\mathbf{X}$ *in the sense that*

$$\lim_{N\to\infty} \overline{\lim_{\lambda\to\infty}}\, \mathcal{P}\{\lambda^{-2}\max(|\omega(t)| : 0 \leq t \leq e^\lambda) \geq N\} = 0.$$

It seems that there is no proof of the existence of the limiting distribution of $\lambda^{-2}\omega(e^\lambda)$ as $\lambda \to \infty$. It is to be noted, however, that Mathieu [19] gave the existence proof together with an explicit representation of the limiting distribution of $\lambda^{-2}\omega(e^\lambda)$ in terms of the local time of $|W|$ at level 0 *when* W *is replaced by* $|W|$.

The above results on recurrence and localization rely heavily on the (asymptotic) self-similarity of W as well as the symmetry of $\mathbf{X}_W$. Without these conditions the situation will change much. In the case of random walks there is a profound work by Bricmont and Kupiainen [2].

References

[1] V. I. Afanas'ev, *On a maximum of a transient random walk in random environment*, Theor. Probab. Appl. **35** (1990), 205–215.
[2] J. Bricmont and A. Kupiainen, *Random walks in asymmetric random environments*, Comm. Math. Phys. **142** (1991), 345–420.
[3] T. Brox, *A one-dimensional diffusion process in Wiener medium*, Ann. Probab. 14 (1986), 1206–1218.
[4] R. Durrett, *Multidimensional random walks in random environments with subclassical limiting behavior*, Comm. Math. Phys. **104** (1986), 87–102.
[5] M. Fukushima, *On recurrence criteria in the Dirichlet space theory*, in Local Times to Global Geometry, Control and Physics (ed. Elworthy), Research Notes in Math., Longman Scientific & Technical, Essex **150** (1986), 100–110.
[6] A. O. Golosov, *Localization of random walks in one-dimensional random environments*, Comm. Math. Phys. **92** (1984), 491–506.
[7] A. O. Golosov, *On limiting distributions for a random walk in a critical one-dimensional random environment*, Russian Math. Surveys 41 (1986), 199–200.
[8] K. Ichihara, *Some global properties of symmetric diffusion processes*, Publ. Res. Inst. Math. Sci., Kyoto Univ. 14 (1978), 441–486.
[9] Y. Kasahara, *Spectral theory of generalized second order differential operators and its applications to Markov processes*, Japan. J. Math. (New Ser.) 1 (1975), 67–84.
[10] K. Kawazu, Y. Tamura, and H. Tanaka, *One-dimensional diffusions and random walks in random environments*, Lecture Notes in Math., Springer-Verlag, Berlin and New York, **1299** (1988), 170–184.
[11] K. Kawazu, Y. Tamura and H. Tanaka, *Localization of diffusion processes in one-dimensional random environment*, J. Math. Soc. Japan **44** (1992), 515–550.

[12] K. Kawazu and H. Tanaka, *On the maximum of a diffusion process in a drifted Brownian environment*, Lecture Notes in Math. (Séminaire de Probabilités XXVII), Springer-Verlag, Berlin and New York, **1557** (1993), 78–85.

[13] K. Kawazu and H. Tanaka, *A diffusion process in a Brownian environment with drift*, to appear in J. Math. Soc. Japan.

[14] H. Kesten, *The limit distribution of Sinai's random walk on random environment*, Phys. A **138** (1986), 299–309.

[15] H. Kesten, M. V. Kozlov and F. Spitzer, *A limit law for random walk in a random environment*, Compositio Math. **30** (1975), 145–168.

[16] S. Kotani and S. Watanabe, *Krein's spectral theory of strings and generalized diffusion processes*, Lecture Notes in Math., Springer-Verlag, Berlin and New York, **923** (1982), 235–259.

[17] A. V. Letchikov, *Localization of one-dimensional random walks in random environments*, Soviet Sci. Rev. Sect. C: Math. Phys. Rev. **8** (1989), 173–220.

[18] E. Marinari, G. Parisi, D. Ruelle, and P. Windey, *On the interpretation of* $1/f$ *noise*, Comm. Math. Phys. **89** (1983), 1–12.

[19] P. Mathieu, *Zero white noise limit through Dirichlet forms, with application to diffusions in a random medium*, Probab. Theory Related Fields **99** (1994), 549–580.

[20] J. Nash, *Continuity of solutions of parabolic and elliptic equations*, Amer. J. Math. **80** (1958), 931–953.

[21] Y. Ogura, *One-dimensional bi-generalized diffusion processes*, J. Math. Soc. Japan 41 (1989), 213–242.

[22] S. Schumacher, *Diffusions with random coefficients*, Contemp. Math. 41 (1985), 351–356.

[23] Ya. G. Sinai, *The limiting behavior of a one-dimensional random walk in a random medium*, Theor. Probab. Appl. **27** (1982), 256–268.

[24] F. Solomon, *Random walks in a random environment*, Ann. Probab. **3** (1975), 1–31.

[25] H. Tanaka, *Limit distribution for 1-dimensional diffusion in a reflected Brownian medium*, Lecture Notes in Math. (Séminaire de Probabilités XXI), Springer-Verlag, Berlin and New York, **1247** (1987), 246–261.

[26] H. Tanaka, *Limit theorem for one-dimensional diffusion process in Brownian environment*, Lecture Notes in Math., Springer-Verlag, Berlin and New York, **1322** (1988), 156–172.

[27] H. Tanaka, *Recurrence of a diffusion process in a multidimensional Brownian environment*, Proc. Japan Acad. Ser. A Math. Sci. **69** (1993), 377–381.

[28] H. Tanaka, *Localization of a diffusion process in a one-dimensional Brownian environment*, Comm. Pure Appl. Math. **47** (1994), 755–766.

[29] M. Yor, *On some exponential functionals of Brownian motion*, Adv. Appl. Probab. **24** (1992), 509–531.

Environment-wise central limit theorem for a diffusion in a Brownian environment with large drift

Hiroshi Tanaka

Department of Mathematics, Faculty of Science and Technology, Keio University, Hiyoshi, Yokohama, Kanagawa 223, Japan

Introduction

General theory of one-dimensional diffusion process was established by Itô and McKean ([1]) more than thirty years ago. In this paper we discuss the central limit theorem concerning a diffusion process in one-dimensional Brownian environment with large drift. Many of our methods benefit from the theory by Itô and McKean.

By a diffusion process in a Brownian environment with drift we mean a process $\mathbf{X}^{x_0} = \{\omega(t), t \geq 0, \mathcal{P}^{x_0}\}$ which is defined in the following way.
(i) W is the space of continuous functions on $\mathbb{R}$ vanishing at 0 and is endowed with the Wiener measure P. For an element w of W, w_κ denotes an element of W defined by $w_\kappa(x) = w(x) - \kappa x/2$ where κ is a given positive constant.
(ii) For $w \in W$ and $x_0 \in \mathbb{R}$, $P_w^{x_0}$ denotes the probability measure on $\Omega = C[0,\infty)$ such that $\mathbf{X}_w^{x_0} = \{\omega(t), t \geq 0, P_w^{x_0}\}$ is a diffusion process with generator

$$\mathcal{L}_w = \frac{1}{2} e^{w_\kappa(x)} \frac{d}{dx}\left(e^{-w_\kappa(x)} \frac{d}{dx}\right)$$

starting from x_0; $\omega(t)$ denotes the value of a trajectory $\omega(\in \Omega)$ at time t.
(iii) $\{\omega(t), t \geq 0\}$ is regarded as a process defined on the probability space $(W \times \Omega, \mathcal{P}^{x_0})$ where $\mathcal{P}^{x_0}(dwd\omega) = P(dw)P_w^{x_0}(d\omega)$. This process is what we denote by $\mathbf{X}^{x_0} = \{\omega(t), t \geq 0, \mathcal{P}^{x_0}\}$; x_0 indicates the starting point of our process.

The process $\mathbf{X}^{x_0}$ is a diffusion analogue of what was considered by Kesten-Kozlov-Spitzer [4] and Solomon [6] as a random walk in a random environment. Asymptotic behavior of the process $\mathbf{X}^0$ as t becomes large was investigated by Kawazu-Tanaka [3]. The results are analogous to those of [4][6] and vary with κ. In particular, when $\kappa > 1$, it was proved that

$$\lim_{x\to\infty} T_x/x = m, \quad \lim_{t\to\infty} \omega(t)/t = 1/m, \quad \mathcal{P}^0\text{-a.s.},$$

where $m = 4(\kappa-1)^{-1}$ and T_x is the hitting time: $T_x = \inf\{t > 0 : \omega(t) = x\}$.

The next problem is then to study a fluctuation. In this paper we assume $\kappa > 2$ and discuss central limit theorems concerning T_x and $\omega(t)$. We put

$$M_x = 2\int_0^x dy \int_{-\infty}^y e^{w_\kappa(y)-w_\kappa(z)}dz,$$

$$\mu(t) = \text{ the inverse function of } M_x,$$

$$\overline{\omega}(t) = \max\{\omega(s) : 0 \le s \le t\},\ \underline{\omega}(t) = \inf\{\omega(s) : s \ge t\},$$

$$A = 64(\kappa-1)^{-2}(\kappa-2)^{-1}.$$

Then it will be seen that $E_w^0\{T_x\} = M_x$, $E\{M_x\} = mx$ and $E\{\mathrm{Var}_w^0(T_x)\} = Ax$ for $x \ge 0$ where E_w^0 and Var_w^0 denotes the expectation and variance under P_w^0, respectively.

Our results are the following where $\kappa > 2$ is assumed.

Environment-wise central limit theorem

(I) *For almost all w with respect to P the process*

$$\left\{\frac{T_{\lambda x} - M_{\lambda x}}{\sqrt{A\lambda}}, x \ge 0, P_w^0\right\}$$

converges in law to a Brownian motion as $\lambda \to \infty$ (in the sense of convergence of probability measures on the Skorohod space).

(II) *For almost all w the process*

$$\left\{\frac{\omega(\lambda t) - \mu(\lambda t)}{\sqrt{m^{-3}A\lambda}}, t \ge 0, P_w^0\right\}$$

converges in law to a Brownian motion as $\lambda \to \infty$. The same is true when $\omega(\lambda t)$ is replaced by either of $\overline{\omega}(\lambda t)$ and $\underline{\omega}(\lambda t)$.

Central limit theorem in random environment, namely CLT under $\mathcal{P}^0$ will be discussed in another occasion.

§1. Some preliminaries

1.1. We compute the first and the second moments of various quantities related to hitting times for $\mathbf{X}^0$. Let $T_{a,b} = \inf\{t > 0 : \omega(t) = a \text{ or } b\}, b < a$. Then it is well-known that

$$u(x) = E_w^x\{\exp(-\lambda\int_0^{T_{a,b}} f(\omega(t))dt)\}$$

satisfies (e.g. see [1])

$$\text{(1.1)} \qquad \begin{cases} (\lambda f - \mathcal{L}_w)u = 0 & \text{in } (b,a), \\ u(a) = u(b) = 1. \end{cases}$$

For $n \ge 0$ we put $u_n(x) = E_w^x\{(\int_0^{T_{a,b}} f(\omega(t))dt)^n\}$. Then $u(x) = \sum_{n=0}^\infty u_n(x) \times (-\lambda)^n/n!$ and (1.1) imply $\sum_{n=0}^\infty (\lambda f - \mathcal{L}_w)u_n(-\lambda)^n/n! = 0$. Therefore we have $u_0 \equiv 1$ and for $n \ge 1$

$$(1.2)\qquad \begin{cases} \mathcal{L}_w u_n = -n u_{n-1} f & \text{in } (b,a), \\ u_n(a) = u_n(b) = 0. \end{cases}$$

Let $S_w(x) = \int_0^x \exp\{w_\kappa(y)\}dy$, the canonical scale of $\mathcal{L}_w$. Then

$$(1.3)\qquad S_w(-\infty) = -\infty, \quad P\text{-a.s., if } \kappa > 0.$$

The following formulas (1.4) and (1.5), which hold for almost all w, can be proved under the assumption $\kappa > 0$ by making use of (1.2) and (1.3): For $x < a$,

$$(1.4)\qquad E_w^x\{T_a\} = 2\int_x^a e^{w_\kappa(y)}dy \int_{-\infty}^y e^{-w_\kappa(z)}dz;$$

$$(1.5)\quad E_w^x\{T_a^2\} = 8\int_x^a e^{w_\kappa(y)}dy\int_{-\infty}^y e^{-w_\kappa(z)}dz\int_z^a e^{w_\kappa(u)}du\int_{-\infty}^u e^{-w_\kappa(v)}dv.$$

Making use of the above formulas we can compute $\mathrm{Var}_w^0\{T_x\}$ for $x > 0$; the result is

$$(1.6)\quad \mathrm{Var}_w^0\{T_x\} = 8\int_0^x dy\int_{-\infty}^y e^{w_\kappa(y)-w_\kappa(z)}dz\int_z^y dy\int_{-\infty}^u e^{w_\kappa(u)-w_\kappa(v)}dv.$$

It will be useful for our later discussions to introduce a one-parameter family of measure preserving transformations $\theta_t, t \in \mathbb{R}$, on (W, P) defined by $(\theta_t w)(x) = w(x+t) - w(t), x \in \mathbb{R}$. It is easy to see that $\theta_t\theta_s = \theta_{t+s}$ and $\{\theta_t\}$ is ergodic. If $\kappa > 0$, then

$$(1.7)\qquad f_0(w) = \int_{-\infty}^0 e^{-w_\kappa(t)}dt$$

is finite (P-a.s.) and $\theta_t f_0 \equiv f_0(\theta_t w) = \int_{-\infty}^t e^{w_\kappa(t)-w_\kappa(s)}ds$. From (1.4) and (1.6) we have for $x > 0$

$$(1.8)\qquad E_w^0\{T_x\} = M_x = 2\int_0^x \theta_y f_0 dy,$$

$$\begin{aligned}(1.9)\quad \mathrm{Var}_w^0\{T_x\} &= 8\int_0^x dy\int_{-\infty}^y e^{w_\kappa(y)-w_\kappa(z)}(\theta_z f_0)^2 dz \\ &= 8\int_0^x \theta_y g dy, \quad \left(g(w) = \int_{-\infty}^0 e^{-w_\kappa(t)}(\theta_t f_0)^2 dt\right).\end{aligned}$$

Making use of $E\{\exp(w_\kappa(x) - w_\kappa(y))\} = \exp\{-\gamma(x-y)\}$ for $x > y$ where

$$\gamma = (\kappa - 1)/2,$$

we have $E\{f_0\} = 1/\gamma$ and hence $E\{M_x\} = mx$, $x \ge 0$, if $\kappa > 1$ (note that $E\{f_0\} < \infty$ iff $\kappa > 1$ and $E\{{f_0}^2\} < \infty$ iff $\kappa > 2$). From now on we assume $\kappa > 2$. The constant A given in the introduction is expressed as $A = 16\gamma^{-2}(2\gamma-1)^{-1}$. We put $f = f_0 - \gamma^{-1}$. Then we have the following:

$$E\{{f_0}^2\} = 2\gamma^{-1}(2\gamma-1)^{-1}; \tag{1.10}$$

$$E\{f\theta_x f\} = \gamma^{-2}(2\gamma-1)^{-1}e^{-\gamma x} \quad \text{for } x \geq 0; \tag{1.11}$$

$$E\{\mathrm{Var}_w^0(T_x)\} = Ax \quad \text{for } x \geq 0. \tag{1.12}$$

1.2. The process $\{\theta_x f_0, x \in \mathbb{R}, P\}$ will often play an important role in our discussions. We make use of habitual notation t (instead of x) to indicate time. By an application of Itô's formula we have

$$d\theta_t f_0 = \theta_t f_0 dw(t) - (\gamma\theta_t f_0 - 1)dt, \tag{1.13}$$

and hence $\theta_t f_0$ is a stationary diffusion process obtained as the unique stationary positive solution of the stochastic differential equation. It can be also written as

$$\theta_t f_0 - f_0 = \int_0^t \theta_s f_0 dw(s) - \gamma \int_0^t \theta_s f ds. \tag{1.14}$$

From now on we assume $\kappa > 2$. Then we can easily prove that $E\{{f_0}^{2\delta}\} < \infty$ if $0 \leq \delta < \kappa/2$. Therefore by the Burkholder-Davis-Gundy inequalities we have

$$E\left\{\max_{0\leq s\leq t}\left|\int_0^s \theta_s f_0 dw(s)\right|^{2\delta}\right\} \leq \text{const.}t^\delta, \quad (t \geq 0, 1 \leq \delta < \kappa/2), \tag{1.15}$$

where const. means a constant that may depend on δ but not on t; such const. will also appear in later discussions and may vary from place to place.

Lemma 1. *$t^{-1/2}\max\{\theta_s f_0 : |s| \leq t\} \to 0$ as $t \to \infty$, P-a.s.*

Proof. Since the stationary diffusion $\theta_t f_0$ is reversible, it is enough to prove that $t^{-1/2}\max\{\theta_s f_0 : 0 \leq s \leq t\} \to 0$ as $t \to \infty$, P-a.s. Define stopping times σ_n, $n \geq 0$, by $\sigma_0 = \inf\{t > 0 : \theta_t f_0 = 2\}$ and $\sigma_n =$ the time of first return of $\theta_t f_0$ to 2 visiting 1 after σ_{n-1} $(n \geq 1)$ and then consider the random variables $X_n = \max\{\theta_t f_0 : \sigma_{n-1} \leq t \leq \sigma_n\}, n \geq 1$. Then $X_n, n \geq 1$, are i.i.d. random variables. Since $\sigma_n/n \to \text{const.} > 0$ (a.s.),

$$\begin{aligned}\overline{\lim_{t\to\infty}}\, t^{-1/2}\max_{0\leq s\leq t}\theta_s f_0 &= \overline{\lim_{n\to\infty}}\, {\sigma_n}^{-1/2}\max_{1\leq k\leq n} X_k \\ &= \text{const.}\ \overline{\lim_{n\to\infty}}\, n^{-1/2}\max_{1\leq k\leq n} X_k, \quad \text{a.s.,}\end{aligned}$$

and the rightmost hand of the above equals to 0, a.s., because

$$P\{X_1 > x\} = \{S(2) - S(1)\}\{S(x) - S(1)\}^{-1} \sim \text{const.}x^{-\kappa}, x \to \infty,$$

(here $S(x)$ is the canonical scale of the diffusion $\theta_t f_0$; it is given by $S'(x) = x^{\kappa-1}\exp(2/x), x > 0$). This proves the lemma.

Lemma 2. *For any* $c_1 > 0$

$$M_{t+u} - M_t = mu(1 + o(1)) + o(\sqrt{\lambda}), \quad |t| \le c_1\lambda, \ u \in \mathbb{R},$$

where $o(1)$ *represents a general term that tends to* 0 *as* $\lambda \to \infty$ *uniformly in* (t, u) *such that* $|t| \le c_1\lambda$ *and* $u \in \mathbb{R}$, *for almost all* w; $o(\sqrt{\lambda})$ *is a term that can be expressed as* $o(1)\sqrt{\lambda}$.

Proof. Step1. For any positive constants c_1 and c_2

$$\text{(1.16)} \qquad \sup \left\{ \left| \frac{M_{t+u} - M_t}{u} - m \right| : |t| \le c_1\lambda, |u| \ge c_2\sqrt{\lambda} \right\} \to \quad 0 \quad \text{as } \lambda \to \infty, \ P\text{-a.s.}$$

We give a proof in the case $c_1 = c_2 = 1$ since the case of general c_1 and c_2 can be treated without any essential change in the proof. We use the notation $\sup_{(\lambda)}$ to denote supremum taken over all t, u satisfying $|t| \le \lambda$ and $u \ge \sqrt{\lambda}$. Since $M_t'/2 = \theta_t f_0$ is a stationary reversible process, it is enough to prove that

$$\text{(1.17)} \qquad \sup{}_{(\lambda)} \left| \frac{1}{u} \int_t^{t+u} \theta_s f\, ds \right| \to 0 \text{ as } \lambda \to \infty, \text{ a.s.}$$

By Lemma 1 we have $\sup_{(\lambda)} u^{-1} |\theta_t f_0 - \theta_{t+u} f_0| \to 0$ as $\lambda \to \infty$, a.s. Combining this fact with (1.14) we see that, for the proof of (1.17), it is enough to show that

$$\text{(1.18)} \qquad \sup{}_{(\lambda)} \left| \frac{1}{u} \int_t^{t+u} \theta_s f_0 dw(s) \right| \to 0 \text{ as } \lambda \to \infty, \text{ a.s.}$$

To prove (1.18) we take δ and α such that

$$\text{(1.19)} \qquad 1 < \delta < \kappa/2, \quad 0 < \alpha < 1/2, \quad \delta/2 + \alpha - 1 > 0,$$

and prepare the estimate

$$\text{(1.20)} \qquad E\{R(\lambda)^{2\delta}\} \le \text{const.}\lambda^{-\delta_1},$$

where $\delta_1 = \delta/2 + \alpha - 1$ and

$$R(t, u) = \frac{1}{u} \int_t^{t+u} \theta_s f_0 dw(s), \quad R(\lambda) = \sup{}_{(\lambda)} |R(t, u)|.$$

Putting $I_k = [k\lambda^\alpha, (k+1)\lambda^\alpha]$ for an integer k and denoting by $\sup_{(k,\lambda)}$ the supremum over all $t \in I_k$ and $u \ge \sqrt{\lambda}$, we write

$$\begin{aligned} & \sup{}_{(k,\lambda)} |R(t, u) - R(k\lambda^\alpha, u)| \\ = \ & \sup{}_{(k,\lambda)} \frac{1}{u} \left| -\int_{k\lambda^\alpha}^t \theta_s f_0 dw(s) + \int_{k\lambda^\alpha + u}^{t+u} \theta_s f_0 dw(s) \right| \le U + V, \end{aligned}$$

where

$$U = \lambda^{-1/2} \max_{0\le r\le\lambda^\alpha} \left| \int_{k\lambda^\alpha}^{k\lambda^\alpha+r} \theta_s f_0 dw(s) \right|,$$

$$V = \sup_{u\ge\sqrt{\lambda}} \max_{0\le r\le\lambda^\alpha} \left| \frac{1}{u} \int_{k\lambda^\alpha+u}^{k\lambda^\alpha+u+r} \theta_s f_0 dw(s) \right|.$$

Making use of (1.15) we have

$$E\{U^{2\delta}\} \le \text{const.}\lambda^{-(1-\alpha)\delta}, \tag{1.21}$$

and $E\{V^{2\delta}\}$ is dominated by

$$\sum_{n=1}^{\infty} E\left\{ \sup_{\sqrt{\lambda}n\le u\le\sqrt{\lambda}(n+1)} \max_{0\le r\le\lambda^\alpha} \left| \frac{1}{u} \int_{k\lambda^\alpha+u}^{k\lambda^\alpha+u+r} \theta_s f_0 dw(s) \right|^{2\delta} \right\}$$
$$\le \lambda^{-\delta} \sum_{n=1}^{\infty} n^{-2\delta} E\{R_\lambda^{2\delta}\} = \text{const.}\lambda^{-\delta} E\{R_\lambda^{2\delta}\},$$

where

$$R_\lambda = \sup_{0\le u\le\sqrt{\lambda}} \max_{0\le r\le\lambda^\alpha} \left| \int_u^{u+r} \theta_s f_0 dw(s) \right|.$$

If $0 \le u \le \sqrt{\lambda}$ and $0 \le r \le \lambda^\alpha$, then u and $u+r$ are simultaneously contained in one of the intervals $[l\lambda^\alpha, l\lambda^\alpha + 2\lambda^\alpha]$, $l = 0, 1, \ldots, l_\lambda$, where $l_\lambda = [\lambda^{\frac{1}{2}-\alpha}]$. Therefore

$$R_\lambda \le 2 \max_{0\le l\le l_\lambda} \max_{0\le r\le 2\lambda^\alpha} \left| \int_{l\lambda^\alpha}^{l\lambda^\alpha+r} \theta_s f_0 dw(s) \right|,$$

and hence $E\{V^{2\delta}\}$ is dominated by

$$\text{const.}\lambda^{-\delta}(l_\lambda + 1) E\left\{ \max_{0\le r\le 2\lambda^\alpha} \left| \int_0^r \theta_s f_0 dw(s) \right|^{2\delta} \right\}$$
$$\le \text{const.}\lambda^{-(1-\alpha)(\delta-1)-\frac{1}{2}} \ (\text{use}(1.15)).$$

This combined with (1.21) implies

$$E\left\{ \sup{}_{(k,\lambda)} |R(t,u) - R(k\lambda^\alpha, u)|^{2\delta} \right\} \le \text{const.}\lambda^{-(1-\alpha)(\delta-1)-\frac{1}{2}}, \tag{1.22}$$

because $(1-\alpha)\delta > (1-\alpha)(\delta-1) + \frac{1}{2}$. Let $k(t)$ be the integer determined from t by $k(t)\lambda^\alpha < t \le (k(t)+1)\lambda^\alpha$ and put $K_\lambda = \{k(t) : |t| \le \lambda\}$. Then $\#K_\lambda \le \text{const.}\lambda^{1-\alpha}$ and (1.22) implies

$$E\left\{ \sup{}_{(\lambda)} |R(t,u) - R(k(t)\lambda^\alpha, u)|^{2\delta} \right\} \le \text{const.}\#K_\lambda \cdot \lambda^{-(1-\alpha)(\delta-1)-\frac{1}{2}} \le \text{const.}\lambda^{-\delta_1}, \tag{1.23}$$

because $(1-\alpha)(\delta-2)+\frac{1}{2}>\delta_1$. By a similar method we can prove that

$$E\left\{\sup_{(\lambda)}|R(k(t)\lambda^{\alpha},u)|^{2\delta}\right\}\le \text{const.}\lambda^{-\delta_1} \tag{1.24}$$

and hence (1.20) follows from (1.23) and (1.24).

We finally prove (1.18). If β is a constant such that $\beta\delta_1>1$, then $E\{R(n^\beta)^{2\delta}\}\le \text{const.}n^{-\beta\delta_1}$ by (1.20) so by using the Chebyschev inequality and the Borel-Cantelli lemma we see that $R(n^\beta)\to 0$ as $n\to\infty$, a.s. Likewise

$$R_c(n^\beta)\equiv \sup\left\{|R(t,u)|:|t|\le cn^\beta, u\ge n^{\beta/2}\right\}\to 0,\ \text{a.s.},$$

for any constant $c>0$. The assertion (1.18) now follows from $R(\lambda)\le R_c(n(\lambda)^\beta)$ with $c=2^\beta$ where $n(\lambda)$ is the integer satisfying $n(\lambda)^\beta<\lambda\le (n(\lambda)+1)^\beta$.

Step 2 is to complete the proof of the lemma. The result (1.16) implies that

$$M_{t+u}-M_t=(m+o(1))u,\quad \text{if } |t|\le c_1\lambda \text{ and } |u|\ge\sqrt{\lambda}/2. \tag{1.25}$$

When t and u are restricted to $|t|\le c_1\lambda$ and $|u|<\sqrt{\lambda}/2$, we write $M_{t+u}-M_t=M'+M''$ where

$$M'=M_{t-\sqrt{\lambda}+(\sqrt{\lambda}+u)}-M_{t-\sqrt{\lambda}},\quad M''=M_{t-\sqrt{\lambda}}-M_t.$$

Then applying (1.16) we have $M'=(m+o(1))(\sqrt{\lambda}+u)$ and $M''=-(m+o(1))\sqrt{\lambda}$. Therefore $M_{t+u}-M_t=mu(1+o(1))+o(\sqrt{\lambda})$ if $|t|\le c_1\lambda$ and $|u|<\sqrt{\lambda}/2$. This combined with (1.25) proves the lemma.

§2. Environment-wise central limit theorem

2.1. Proof of (I). We put $\tau_k=T_k-T_{k-1}, k\ge 1$. Then $\tau_k, k\ge 1$, are independent under the probability measure P_w^0. Since $T_n=\sum_1^n\tau_k$, the central limit theorem for T_n with fixed w can be obtained by verifying the Lindeberg condition. We have

$$\text{Var}_w^0\{T_n\}=\sum_{k=1}^n\text{Var}_w^0\{\tau_k\}=\sum_{k=1}^n V(\theta_{k-1}w)$$

where $V(w)=\text{Var}_w^0\{\tau_1\}$. We now assume $\kappa>2$ to ensure the existence of $E\{V\}$, which is equal to A by (1.3). Since $\{\theta_t\}$ is ergodic, we have

$$\lim_{n\to\infty}\frac{\text{Var}_w^0\{T_n\}}{n}=A,\ P\text{-a.s.} \tag{2.1}$$

Putting $\tilde{\tau}_k = \tau_k - E_w^0\{\tau_k\}$ and $V_N(w) = E_w^0\{|\tilde{\tau}_1|^2; |\tilde{\tau}_1| > N\}$ (the integral of $|\tilde{\tau}_1|^2$ over $\{|\tilde{\tau}_1| > N\}$ with respect to P_w^0), we have for any $\varepsilon > 0$ and for almost all w

$$(2.2)\qquad \begin{aligned} &\frac{1}{\mathrm{Var}_w^0\{T_n\}} \sum_{k=1}^n E_w^0 \left\{ |\tilde{\tau}_k|^2; |\tilde{\tau}_k| > \varepsilon\sqrt{\mathrm{Var}_w^0\{T_n\}} \right\} \\ \le\ & \frac{\text{const.}}{n} \sum_{k=1}^n E_w^0 \left\{ |\tilde{\tau}_k|^2; |\tilde{\tau}_k| > N \right\} \ \text{(by (2.1))} \\ =\ & \frac{\text{const.}}{n} \sum_{k=1}^n V_N(\theta_{k-1}w) \to \text{const.} E\{V_N\}, \quad n \to \infty. \end{aligned}$$

Since $E\{V_N\} \downarrow 0$ as $N \uparrow \infty$ because of $E\{V\} < \infty$, the left hand side of (2.2) tends to 0 as $n \uparrow \infty$ (P-a.s.), which shows that the Lindeberg condition is satisfied for $\{\tilde{\tau}_k\}$. Therefore for almost all w the central limit theorem holds for T_n with respect to P_w^0. By (1.4) $E_w^0\{T_n\} = M_n$ and by (2.1) $\mathrm{Var}_w^0\{T_n\} \sim An, n \to \infty$ (P-a.s.). Therefore the Lindeberg condition together with an application of Theorem 3.1 of [5] implies the assertion of (I). Further details will be given in a joint paper with K. Kawazu.

2.2. Proof of (II). The probability measure we consider in this proof is P_w^0 where w is taken from some subset of W that has P-measure 1. Thus events we consider here can be regarded as subsets of $\Omega_0 = \{\omega \in \Omega : \omega(0) = 0\}$. For any $s > 0$, $\rho \in \mathbb{R}$ and $\nu > 0$ we have, with convention $\underline{\omega}(\infty) = \infty$,

$$(2.3)\qquad \begin{aligned} \{T_{\rho^+} < s\} &= \{\overline{\omega}(s) > \rho\} \supset \{\omega(s) > \rho\} \supset \{\underline{\omega}(s) > \rho\} \\ &\supset \{T_{(\rho+\nu)^+} < s\} \cap \{\underline{\omega}(T_{\rho+\nu}) > \rho\}, \end{aligned}$$

where $\rho^+ = \max\{\rho, 0\}$ and T_x denotes the hitting time to $x \in \mathbb{R}$. Similarly

$$(2.4)\qquad \begin{aligned} \{T_{\rho^+} > s\} &= \{\overline{\omega}(s) < \rho\} \subset \{\omega(s) < \rho\} \subset \{\underline{\omega}(s) < \rho\} \\ &\subset \{T_{(\rho+\nu)^+} > s\} \cup \{\underline{\omega}(T_{\rho+\nu}) < \rho\}. \end{aligned}$$

Lemma 3. *For any fixed $t_1 > 0$ let (φ, ψ) be a pair of continuous functions on $[0, t_1]$ such that*

$$(2.5)\qquad \begin{aligned} &\varphi \textit{ and } \psi \textit{ are absolutely continuous, } \varphi', \psi' \in L^2[0, t_1], \\ &\varphi(t) < \psi(t) \textit{ for all } t \in [0, t_1] \textit{ and } \varphi(0) \neq 0, \psi(0) \neq 0. \end{aligned}$$

Then for almost all w

$$(2.6)\qquad \begin{aligned} \lim_{\lambda\to\infty}\ & P_w^0 \left\{ \varphi(t) < \frac{\omega(\lambda t) - \mu(\lambda t)}{\sqrt{m^{-3}A\lambda}} < \psi(t) \textit{ for all } t \in [t_0, t_1] \right\} \\ =\ & P\{\varphi(t) < w(t) < \psi(t) \textit{ for all } t \in [t_0, t_1]\}. \end{aligned}$$

Proof. For the proof it is enough to consider the case $\varphi(0) < 0 < \psi(0)$. We put $\rho(t) = \mu(\lambda t) + \varphi(t)\sqrt{m^{-3}A\lambda}$ and $\rho_1(t) = \mu(\lambda t) + \psi(t)\sqrt{m^{-3}A\lambda}$. Then $\rho_1(t) > 0$ for all $t \in [0,t_1]$ if λ is sufficiently large, but $\rho(t) < 0$ if t is sufficiently small for each fixed λ. Since $\{\varphi(t) < (m^{-3}A\lambda)^{-1/2}(\omega(\lambda t) - \mu(\lambda t)) < \psi(t)\} = \{\rho(t) < \omega(\lambda t) < \rho_1(t)\}$, an application of (2.3) and (2.4) yields

$$\begin{aligned}
&\{T_{(\rho(t)+\nu)^+} < \lambda t, T_{\rho_1(t)} > \lambda t, \underline{\omega}(T_{\rho(t)+\nu}) > \rho(t)\}\\
&\subset \{\varphi(t) < (m^{-3}A\lambda)^{-1/2}(\omega(\lambda t) - \mu(\lambda t)) < \psi(t)\}\\
&\subset \{T_{\rho(t)^+} < \lambda t\} \cap [\{T_{\rho_1(t)+\nu} > \lambda t\} \cup \{\underline{\omega}(T_{\rho_1(t)+\nu}) < \rho_1(t)\}].
\end{aligned}$$

Therefore if we put

$$\begin{aligned}
\Gamma_\lambda &= \{\varphi(t) < (m^{-3}A\lambda)^{-1/2}(\omega(\lambda t) - \mu(\lambda t)) < \psi(t) \text{ for all } t \in [0, t_1]\},\\
\Gamma_\lambda^- &= \{T_{(\rho(t)+\nu)^+} < \lambda t, T_{\rho_1(t)} > \lambda t \text{ for all } t \in [0,t_1]\},\\
\Gamma_\lambda^+ &= \{T_{\rho(t)^+} < \lambda t, T_{\rho_1(t)+\nu} > \lambda t \text{ for all } t \in [0,t_1]\},\\
A_\lambda &= \{\underline{\omega}(T_{\rho(t)+\nu}) > \rho(t) \text{ for all } t \in [0,t_1]\},\\
B_\lambda &= \{\underline{\omega}(T_{\rho_1(t)+\nu}) < \rho_1(t) \text{ for some } t \in [0,t_1]\},
\end{aligned}$$

then

$$\Gamma_\lambda^- \cap A_\lambda \subset \Gamma_\lambda \subset \Gamma_\lambda^+ \cup B_\lambda. \tag{2.7}$$

Put $\nu = 6\log\lambda$ and let us prove

$$\lim_{\lambda\to\infty} P_w^0\{A_\lambda^c\} = \lim_{\lambda\to\infty} P_w^0\{B_\lambda\} = 0, \ P\text{-a.s.} \tag{2.8}$$

Since $\{\underline{\omega}(T_{\rho(t)+\nu}) \le \rho(t)\} \subset \{\underline{\omega}(T_{[\rho(t)+\nu]}) < [\rho(t) + \nu] - (\nu - 1)\}$, we have

$$A_\lambda^c \subset \bigcup_{k=\rho_\lambda^-}^{\rho_\lambda^+} \{\underline{\omega}(T_k) < k - (\nu - 1)\},$$

where $\rho_\lambda^- = \min\{[\rho(t)+\nu] : 0 \le t \le t_1\}$ and $\rho_\lambda^+ = \max\{[\rho(t)+\nu] : 0 \le t \le t_1\}$. Therefore

$$P_w^0\{A_\lambda^c\} \le \sum_{k=\rho_\lambda^-}^{\rho_\lambda^+} P_w^0\{\underline{\omega}(T_k) < k - (\nu - 1)\}.$$

By an ergodic theorem $M_x/x \to m$ as $x \to \infty$ (P-a.s.) and hence $\mu(t)/t \to m^{-1}$ as $t \to \infty$ (P-a.s.), which implies $\rho_\lambda^+ \sim \lambda m^{-1}t_1$ and $\rho_\lambda^- = o(\lambda)$ as $\lambda \to \infty$ (P-a.s.). Therefore for the proof of (2.8) it is enough to show that for any constant $c > 0$

$$\lim_{\lambda\to\infty} \sum_{|k|<c\lambda} P_w^0\{\underline{\omega}(T_k) < k - (\nu - 1)\} = 0, \quad P\text{-a.s.}$$

Taking the expectation we have

$$\begin{aligned}
&E\left\{\sum_{|k|<c\lambda} P_w^0\left\{\underline{\omega}(T_k) < k-(\nu-1)\right\}\right\}\\
&= \sum_{|k|<c\lambda} \mathcal{P}^0\left\{\underline{\omega}(0) < -(\nu-1)\right\}\\
&< \text{const.}\lambda \exp\left\{-(\kappa-1)(\nu-1)/2\right\} \le \text{const.}\lambda^{-2},
\end{aligned}$$

where we used the result of [2]. The above estimate holds for $\nu = 6\log\lambda + O(1)$. Therefore an application of Borel-Cantelli lemma implies

$$\lim_{n\to\infty} \sum_{|k|<cn} P_w^0\left\{\underline{\omega}(T_k) < k-(\nu_n-1)\right\} = 0, \quad P\text{-a.s.},$$

for any given sequence $\{\nu_n\}$ such that $\nu_n = 6\log n + O(1)$ as $n \to \infty$. We thus have

$$\begin{aligned}
&\sum_{|k|<c\lambda} P_w^0\left\{\underline{\omega}(T_k) < k-(\nu-1)\right\}\\
&\le \sum_{|k|<c([\lambda]+1)} P_w^0\left\{\underline{\omega}(T_k) < k-(\nu-1)\right\}\\
&\to 0 \quad \text{as } \lambda \to 0 \ (P\text{-a.s.}),
\end{aligned}$$

which implies the assertion (2.8) concerning A_λ^c. The assertion for B_λ can be proved similarly.

To proceed let $\xi = \lambda m^{-1}$ and put

$$\alpha(t) = \xi^{-1}\rho(t) = \lambda^{-1}m\left\{\mu(\lambda t) + \varphi(t)\sqrt{m^{-3}A\lambda}\right\},$$

$$\beta(t) = \xi^{-1}\rho_1(t) = \lambda^{-1}m\left\{\mu(\lambda t) + \psi(t)\sqrt{m^{-3}A\lambda}\right\}.$$

Then for almost all w both $\alpha(t)$ and $\beta(t)$ tend to t as $\lambda \to \infty$, the convergence being uniform on each finite t-interval. From now on we write $T(x)$ and $M(x)$ for T_x and M_x, respectively. Since $\mu(\lambda t) = \xi\alpha(t) - \varphi(t)\sqrt{m^{-3}A\lambda}$, an application of Lemma 2 yields

$$\text{(2.9)}\quad \lambda t = M(\xi\alpha(t) - \varphi(t)\sqrt{m^{-3}A\lambda}) = M(\xi\alpha(t)) - \varphi(t)\sqrt{m^{-1}A\lambda} + o(\sqrt{\lambda}),$$

where $o(\sqrt{\lambda})$ is a term that, when divided by $\sqrt{\lambda}$, tends to 0 uniformly in $t \in [0, t_1]$ as $\lambda \to \infty$ for almost all w. Similarly

$$\text{(2.10)}\qquad \lambda t = M(\xi\beta(t)) - \psi(t)\sqrt{m^{-1}A\lambda} + o(\sqrt{\lambda}).$$

In what follows the notation $o(1)$ represents a term that tends to 0 uniformly in $t \in [0, t_1]$ as $\lambda \to \infty$ for almost all w. Using (2.9) and (2.10) we can obtain the following (2.11) $\sim$ (2.14) that hold for almost all w:

$$(2.11)\quad T(\rho(t)^+) < \lambda t \Leftrightarrow \frac{T(\xi\alpha(t)^+) - M(\xi\alpha(t)^+)}{\sqrt{A\xi}} < -\varphi(t) + o(1);$$

$$(2.12)\quad T(\rho_1(t)) > \lambda t \Leftrightarrow \frac{T(\xi\beta(t)) - M(\xi\beta(t))}{\sqrt{A\xi}} > -\psi(t) + o(1);$$

$$\begin{aligned}(2.13)\ T\left((\rho(t)+\nu)^+\right) &< \lambda t \Leftrightarrow T\left((\rho(t)+\nu)^+\right) - M(\xi\alpha(t)) \\ &\qquad < -\varphi(t)\sqrt{m^{-1}A\lambda} + o(\sqrt{\lambda}) \\ &\Leftrightarrow \frac{T\left(\xi(\alpha(t)+\frac{\nu}{\xi})^+\right) - M\left(\xi(\alpha(t)+\frac{\nu}{\xi})^+\right)}{\sqrt{A\xi}} \\ &\qquad < -\varphi(t) + o(1);\end{aligned}$$

$$\begin{aligned}(2.14)\qquad & T\left((\rho_1(t)+\nu)\right) > \lambda t \Leftrightarrow \\ & \frac{T\left(\xi(\beta(t)+\frac{\nu}{\xi})\right) - M\left(\xi(\beta(t)+\frac{\nu}{\xi})\right)}{\sqrt{A\xi}} > -\psi(t) + o(1).\end{aligned}$$

In deriving the second equivalence in (2.13) as well as the equivalence (2.14) we again used Lemma 2. From (2.11) $\sim$ (2.14) we have the following: For *almost all* w, $P_w^0\{\Gamma_\lambda^+\}$ is equal to $p_w(\xi)$ where $p_w(\xi)$ is the probability, evaluated by P_w^0, of the event

$$\left\{\begin{array}{l}(A\xi)^{-1/2}\left\{T(\xi\alpha(t)^+) - M(\xi\alpha(t))^+\right\} < -\varphi(t) + o(1), \\ (A\xi)^{-1/2}\left\{T(\xi(\beta(t)+\nu\xi^{-1})) - M(\xi(\beta(t)+\nu\xi^{-1}))\right\} > -\psi(t) + o(1), \\ \qquad\qquad \text{for all } t \in [0,t_1]\end{array}\right\}.$$

But by (I), for almost all w, $p_w(\xi)$ tends to

$$\begin{aligned} & P\{-\psi(t) < w(t) < -\varphi(t) \text{ for all } t \in [0,t_1]\} \\ = \ & P\{\varphi(t) < w(t) < \psi(t) \text{ for all } t \in [0,t_1]\}\end{aligned}$$

as $\xi \to \infty$ (here we used the assumption (2.5)). We also have a similar statement for $P_w^0\{\Gamma_\lambda^-\}$. Therefore from (2.7) and (2.8) we finally obtain Lemma 3. It is also to be noted that, by virtue of (2.3) and (2.4), Lemma 3 remains valid when $\omega(\cdot)$ is replaced by either of $\overline{\omega}(\cdot)$ and $\underline{\omega}(\cdot)$.

For fixed $t_1 > 0$ we denote by $\mathbb{F}$ the set of all pairs (φ, ψ) of functions on $[0, t_1]$ of the form $\varphi = p_1 \vee p_2 \vee \cdots \vee p_n$(maximum) and $\psi = q_1 \wedge q_2 \wedge \cdots \wedge q_n$ (minimum) where p_j and q_j are polynomials with rational coefficients satisfying $p_j(t) < q_j(t)$ for all $t \in [0,t_1], p_j(0) \neq 0, q_j(0) \neq 0 (1 \leq j \leq n)$. Since $\mathbb{F}$ is countable, Lemma 3 implies that for almost all w (2.6) holds for all $(\varphi,\psi) \in \mathbb{F}$. Fixing an arbitrary w for which (2.6) holds for all $(\varphi,\psi) \in \mathbb{F}$, we denote by Q_λ the probability law of the process $\{(m^{-3}A\lambda)^{-1/2}(\omega(\lambda t) - \mu(\lambda t)), 0 \leq t \leq t_1, P_w^0\}$; we also denote by Q the Wiener measure on $C[0,t_1]$ with $Q\{\tilde{w}(0) = 0\} = 1$. A Borel set in the Banach space $C[0,t_1]$ is said to be admissible if $Q_\lambda(A) \to Q(A)$ as $\lambda \to \infty$. Then for each pair $(\varphi, \psi) \in \mathbb{F}$ the open set $U(\varphi,\psi) \equiv \{\tilde{w} \in C[0,t_1] : \varphi(t) < \tilde{w}(t) < \psi(t)$ for all $t \in [0,t_1]\}$ is admissible. On the other hand any open set G in $C[0,t_1]$ can be expressed as $\cup_{n=1}^{\infty} U_n$ where each U_n is of the form $U(\varphi,\psi)$ with $(\varphi,\psi) \in \mathbb{F}$. The intersection of any finite number of U_k's is admissible because it is still of the form $U(\varphi,\psi)$, so $G_n \equiv \cup_{k=1}^{n} U_k$ is admissible because of the inclusion-exclusion formula

$$Q_\lambda(G_n) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \le i_1 < \cdots < i_k \le n} Q_\lambda(U_{i_1} \cap U_{i_2} \cap \cdots \cap U_{i_k}).$$

Since $G_n \uparrow G$ we have $\underline{\lim}_{\lambda\to\infty} Q_\lambda(G) \ge Q(G)$, which implies that $Q_\lambda \to Q$ as $\lambda \to \infty$. Since $t_1 > 0$ is arbitrary, this completes the proof of (II).

References

[1] Itô, K., and McKean, H.P., Diffusion Processes and Their Sample Paths, Springer-Verlag, New York, 1965.

[2] Kawazu, K., and Tanaka,H., On the maximum of a diffusion process in a drifted Brownian environment, Séminaire de Probabilités, LMN 1557, 78–85.

[3] Kawazu, K., and Tanaka,H., A diffusion process in a Brownian environment with drift, to appear in J.Math.Soc.Japan.

[4] Kesten, H., Kozlov, M. V., and Spitzer, F., A limit law for random walk in a random environment, Composito Math. **30** (1975), 145–168.

[5] Prohorov, Yu. V., Convergence of random processes and limit theorems in probability theory, Theor. Probab. Appl. 1 (1956), 157–214.

[6] Solomon, F., Random walks in a random environment, Ann. Probab. **3** (1975),1–31.

Reprinted from

Itô's Stochastic Calculus and Probability Theory,

373–384, Springer-Verlag, 1996.

J. Math. Soc. Japan
Vol. 49, No. 2, 1997

A diffusion process in a Brownian environment with drift*

By Kiyoshi KAWAZU and Hiroshi TANAKA

(Received Feb. 8, 1995)

Introduction.

Let $\boldsymbol{W}$ be the space of continuous functions W defined in $\boldsymbol{R}$ and vanishing at the origin. Let P be the Wiener measure on $\boldsymbol{W}$, namely, the probability measure on $\boldsymbol{W}$ such that $\{W(t), t\geqq 0, P\}$ and $\{W(-t), t\geqq 0, P\}$ are independent one-dimensional Brownian motions. Let $\Omega=C[0, \infty)$ and denote by $\omega(t)$ the value of a function ω $(\in\Omega)$ at time t. Given a sample function W $(\in\boldsymbol{W})$ and a nonnegative constant κ we consider a probability measure P_W^x on Ω such that $\{\omega(t), t\geqq 0, P_W^x\}$ is a diffusion process with generator

$$(1)\qquad \mathcal{L}_W=\frac{1}{2}e^{W_\kappa(x)}\frac{d}{dx}\left(e^{-W_\kappa(x)}\frac{d}{dx}\right)=\frac{dx}{m_W(dx)}\cdot\frac{dx}{dS_W(x)}$$

starting from x, where

$$(2)\qquad W_\kappa(x)=W(x)-\frac{1}{2}\kappa x,$$

$$(3)\qquad S_W(x)=\int_0^x e^{W_\kappa(y)}dy, \quad m_W(dx)=2e^{-W_\kappa(x)}dx.$$

It is well-known that a version of $\{\omega(t), t\geqq 0, P_W^x\}$ can be constructed from a Brownian motion by a scale-change and a time-change. When W is considered *random*, $\{\omega(t), t\geqq 0\}$ is regarded as a process defined on the probability space $(\boldsymbol{W}\times\Omega, \mathcal{P}^x)$ where $\mathcal{P}^x(dWd\omega)=P(dW)P_W^x(d\omega)$. We thus have a process $\boldsymbol{X}^x=\{\omega(t), t\geqq 0, \mathcal{P}^x\}$ which, in this paper, is called a *diffusion process in a Brownian environment with drift*. The following intuitive description may suggest the name. The process $\boldsymbol{X}^x$ is obtained as a formal solution of the symbolic equation

$$(4)\qquad dX(t)=dB(t)-\frac{1}{2}W_\kappa'(X(t))dt,$$

* This work was partially supported by Grant-in Aid for Science (No. 07454034), Ministry of Education of Japan.

where $B(t)$ is a Brownian motion independent of $W(\cdot)$ (however, note that (4) has no rigorous meaning).

When $\kappa>0$ X^0 may be regarded as a diffusion model of a random walk in a random environment discussed by Kozlov [14], Solomon [18] and Kesten-Kozlov-Spitzer [10]. When $\kappa=0$ X^0 is a diffusion model of Sinai's random walk in a random environment ([17]). The asymptotic behavior of $\{\omega(t), \mathcal{P}^0\}$ as $t\to\infty$ when $\kappa=0$ was discussed by Schumacher [16] and Brox [1]. They showed that $\{\omega(t), \mathcal{P}^0\}$ exhibits the same asymptotic behavior as Sinai's random walk, namely, that the limiting distribution of $(\log t)^{-2}\omega(t)$ as $t\to\infty$ exists (see also [2], [3], [8], [11], [19] for related works). When $\kappa>0$ (in particular when $0<\kappa<1$) it was an open problem to obtain results for X^0 which are (or at least expected to be) similar to those of Kesten-Kozlov-Spitzer [10]. The purpose of the present paper is to give some answer to this problem.

Let $T_x=\inf\{t>0:\omega(t)=x\}$, $\bar{\omega}(t)=\max\{\omega(s):0\leqq s\leqq t\}$ and $\underline{\omega}(t)=\inf\{\omega(s):s\geqq t\}$. Then our result in the case $\kappa>0$ is the following.

THEOREM 1. (i) *If* $0<\kappa<1$, *then*

$$\lim_{x\to\infty}\mathcal{P}^0\{x^{-1/\kappa}T_x\leqq t\}=F_\kappa(t),\quad t>0, \tag{5}$$

$$\lim_{t\to\infty}\mathcal{P}^0\{t^{-\kappa}\bar{\omega}(t)\leqq x\}=\lim_{t\to\infty}\mathcal{P}^0\{t^{-\kappa}\omega(t)\leqq x\} \tag{6}$$
$$=\lim_{t\to\infty}\mathcal{P}^0\{t^{-\kappa}\underline{\omega}(t)\leqq x\}=1-F_\kappa(x^{-1/\kappa}),\quad x>0,$$

where F_κ *is the distribution function of a one-sided stable distribution with Laplace transform* $\exp(-c\lambda^\kappa)$; *the constant* c *is given by*

$$c=\left\{2^{1-\kappa}\Gamma(\kappa)\int_0^\infty\frac{dx}{u(x)^2}\right\}^{-1}$$

where $u(x)$ *is the solution of*

$$\frac{d}{dM(x)}\frac{d}{dx}u=2u,\quad u(0)=1,\quad u'(0)=0,$$

the function $M(x)$ *being given in Lemma* 1.

(ii) *If* $\kappa=1$, *then*

(7) $(x\log x)^{-1}T_x$ *converges to* 4 *in probability with respect to* $\mathcal{P}^0$ *as* $x\to\infty$;

(8) *each of* $t^{-1}(\log t)\bar{\omega}(t)$, $t^{-1}(\log t)\omega(t)$ *and* $t^{-1}(\log t)\underline{\omega}(t)$ *converges to* 1/4 *in probability with respect to* $\mathcal{P}^0$ *as* $t\to\infty$.

(iii) *If* $\kappa>1$, *then*

$$\lim_{x\to\infty}T_x/x=\frac{4}{\kappa-1},\quad \mathcal{P}^0\text{-}a.s., \tag{9}$$

$$(10)\qquad \lim_{t\to\infty}\omega(t)/t=\frac{\kappa-1}{4},\quad \mathcal{P}^0\text{-a.s..}$$

The assertion (i) can be slightly strengthened as follows.

THEOREM 2. *Let* $0<\kappa<1$.

(i) *The process* $\{\lambda^{-1/\kappa}T_{\lambda x},\ x\geqq 0,\ \mathcal{P}^0\}$ *converges to* $\{L(x),\ x\geqq 0\}$ *as* $\lambda\to\infty$ *in the sense of convergence of finite dimensional distributions, where* $\{L(x),\ x\geqq 0\}$ *is an increasing stable process with Laplace transform*

$$E\{\exp(-\xi L(1))\}=\exp(-c\xi^{\kappa}),\quad \xi\geqq 0.$$

(ii) *The process* $\{\lambda^{-\kappa}\omega(\lambda t),\ t\geqq 0,\ \mathcal{P}^0\}$ *converges to* $\{L^{-1}(t),\ t\geqq 0\}$ *as* $\lambda\to\infty$ *in the sense of convergence of finite dimensional distributions, where*

$$L^{-1}(t)=\inf\{x>0: L(x)>t\}.$$

In the case of random walks, results similar to (5) and (6) were obtained by Kesten-Kozlov-Spitzer [**10**] and results similar to (7), (8), (9) and (10) by Solomon [**18**]. Our method of proving (9) and (10) is similar to that of [**18**] but as for (5), (6), (7) and (8) our method is different from either of [**10**] and [**18**]; it is based on Kotani's formula (see § 1) which reduces our problem to the study of limiting behavior of another diffusion process described by a certain stochastic differential equation with non-random coefficients. In proving (5) and (6) we must also use Kasahara's continuity theorem ([**7**]) concerning Krein's correspondence ([**6**]) between the m-measure and the spectral measure (or more precisely the h-function) of a one-dimensional diffusion operator.

ACKNOWLEDGMENT. The authors wish to thank S. Kotani for valuable discussions on the subject and for permitting us to contain his formula (1.1) here.

§ 1. Kotani's formula.

The following formula was obtained by S. Kotani in 1988 in his study of the limiting distribution of $(\log t)^{-2}\overline{\omega}(t)$ in the case $\kappa=0$ (unpublished).

KOTANI'S FORMULA. *Let* $\lambda>0$. *Then for* $t\geqq 0$

$$(1.1)\qquad E_W^0\{e^{-\lambda T_t}\}=\exp\left\{-\int_0^t U_\lambda(s)ds\right\},\quad P\text{-a.s.,}$$

where $U_\lambda(t)$ *is the unique stationary positive solution of*

$$(1.2)\qquad dU_\lambda(t)=U_\lambda(t)dW(t)+\left\{2\lambda+\frac{1-\kappa}{2}U_\lambda(t)-U_\lambda(t)^2\right\}dt.$$

PROOF. Taking an arbitrary but fixed $a>0$ we put for $0<t<a$

$$u(t)=1/E_W^0\{e^{-\lambda T_t}\}, \quad v(t)=E_W^t\{e^{-\lambda T_a}\}.$$

Then

$$E_W^0\{e^{-\lambda T_a}\}=E_W^0\{e^{-\lambda T_t}\}E_W^t\{e^{-\lambda T_a}\}$$
$$=v(t)/u(t), \quad P\text{-a.s..}$$

Since $\mathcal{L}_W v(t)=\lambda v(t)$, $t<a$, $u(t)$ also satisfies $\mathcal{L}_W u(t)=\lambda u(t)$, $t>0$, or equivalently

$$d\{e^{-W_\kappa(t)}u'(t)\}=2\lambda e^{-W_\kappa(t)}u(t)dt, \quad t>0. \tag{1.3}$$

If we put $U_\lambda(t)=\{\log u(t)\}'=u'(t)/u(t)$, then $U_\lambda(t)>0$, P-a.s.. Since $u(t)$ and $U_\lambda(t)$ are adapted to the filtration generated by $\{W(t)\}$, we can apply Itô's formula to compute the stochastic differential $dU_\lambda(t)$. Using (1.3) we have

$$\begin{aligned}
dU_\lambda(t) &= d(u'(t)u(t)^{-1})\\
&= d(e^{-W_\kappa(t)}u'(t)e^{W_\kappa(t)}u(t)^{-1})\\
&= e^{W_\kappa(t)}u(t)^{-1}d(e^{-W_\kappa(t)}u'(t))\\
&\quad + e^{-W_\kappa(t)}u'(t)u(t)^{-1}de^{W_\kappa(t)} - e^{-W_\kappa(t)}u'(t)e^{W_\kappa(t)}u(t)^{-2}du(t)\\
&= 2\lambda dt + u'(t)u(t)^{-1}dW_\kappa(t) + 2^{-1}u'(t)u(t)^{-1}dt - \{u'(t)u(t)^{-1}\}^2dt\\
&= U_\lambda(t)dW(t) + \left\{2\lambda + \frac{1-\kappa}{2}U_\lambda(t) - U_\lambda(t)^2\right\}dt.
\end{aligned}$$

For $h>0$ we can write $u(t+h)=u(t)\tilde{u}(h)$ where

$$\tilde{u}(h)=1/E_W^t\{e^{-\lambda T_{t+h}}\} \overset{d}{=} u(h);$$

in the above "$\overset{d}{=}$" means the equality in distribution. Therefore

$$\begin{aligned}
U_\lambda(t) = u'(t)u(t)^{-1} &= \lim_{h\downarrow 0}\frac{\tilde{u}(h)-1}{h}\\
&\overset{d}{=} \lim_{h\downarrow 0}\frac{u(h)-1}{h} = u'(0)\\
&= u'(0)u(0)^{-1} = U_\lambda(0).
\end{aligned}$$

This implies that $U_\lambda(t)$ is a stationary solution of (1.2). The uniqueness of such a solution follows from Theorem 18 of Itô-Nisio [**5**].

§ 2. Kasahara's continuity theorem for Krein's correspondence.

Krein's theory of strings ([**6**]) has many applications to diffusion processes (e.g., see [**12**], [**13**], [**20**]). In this section we do not give the general theory but list some of the results of Kasahara [**7**] on Krein's correspondence that will be useful for our later discussions. For a general statement of Krein's correspondence theory it is convenient to consider inextensible measures (e.g., see [**20**]). From the view point of its application to the present paper, however, it is enough to consider simply Radon measures in $[0, \infty)$. Thus suppose we are given a Radon measure $m(dx)$ in $[0, \infty)$. We exclude the trivial case where $m(dx)=0$. The associated function $M(x)$ is defined by $M(x)=m([0, x))$ for $x>0$ and $M(0)=0$. Consider the generalized differential operator $\mathcal{L}=d/m(dx)\cdot d/dx$ and let $\varphi(x, \alpha)$ and $\psi(x, \alpha)$ be the solutions of $\mathcal{L}u=\alpha u$ with the initial conditions

(2.1) $$u(0)=1, \quad u'(0)=0,$$

(2.2) $$u(0)=0, \quad u'(0)=1,$$

respectively. For $x>0$ $\varphi(x, \alpha)$ and $\psi(x, \alpha)$ satisfy

(2.3) $$\varphi(x, \alpha)=1+\alpha\iint_{0\le z<y<x}\varphi(z, \alpha)m(dz)dy,$$

(2.4) $$\psi(x, \alpha)=x+\alpha\iint_{0\le z<y<x}\psi(z, \alpha)m(dz)dy.$$

The pair $\{\varphi(x, \alpha), \psi(x, \alpha)\}$ is called the system of fundamental solutions associated with $m(dx)$. It is known that for $\alpha>0$

(2.5) $$h(\alpha)=\lim_{x\to\infty}\psi(x, \alpha)/\varphi(x, \alpha)=\int_0^\infty\frac{dx}{\varphi(x, \alpha)^2}<\infty$$

exists. The function $h(\alpha)$ is called the characteristic function of $m(dx)$ or of $M(x)$. The following (2.6), (2.7) and (2.8) are also known.

(2.6) The correspondence between $m(dx)$ and $h(\alpha)$ is one to one ([**6**], see also [**13**]).

(2.7) $a^{-1}h(c\alpha)$ is the characteristic function of $acM(ax)$ for arbitrary positive constants a and c ([**7**]).

(2.8) Assume that ∞ is not regular for $\mathcal{L}$, namely, that at least one of the integrals $\iint_{0<y<x<\infty}m(dy)dx$, $\iint_{0<y<x<\infty}dy\, m(dx)$ diverges. Then for each $\alpha>0$ a positive decreasing solution u of $\mathcal{L}u=\alpha u$ with $u(0)=1$ is unique and expressed as $u(x)=\varphi(x, \alpha)-\psi(x, \alpha)/h(\alpha)$ ([**4**]).

We now state

KASAHARA'S CONTINUITY THEOREM ([7]). *Let $m_n(dx)$, $n=0, 1, \cdots$, be Radon measures in $[0, \infty)$ with associated functions $M_n(x)$, characteristic functions $h_n(\alpha)$ and systems of fundamental solutions $\{\varphi_n(x, \alpha), \psi_n(x, \alpha)\}$. Then the following statements are equivalent to each other.*

(i) *$M_n(x) \to M_0(x)$ as $n \to \infty$ at each continuity point x of $M_0(\cdot)$.*

(ii) *For $x \geqq 0$ and $\alpha > 0$ $\varphi_n(x, \alpha) \to \varphi_0(x, \alpha)$ as $n \to \infty$.*

(iii) *For $\alpha > 0$ $h_n(\alpha) \to h(\alpha)$ as $n \to \infty$.*

§ 3. Proof of Theorem 1 in the case $0<\kappa<1$.

Let $U_\lambda(t)$ be the diffusion process appearing in Kotani's formula and put

$$V_\lambda(t) = \frac{1}{2\lambda} U_\lambda(t).$$

Since $V_\lambda(t)$ satisfies the stochastic differential equation

$$dV_\lambda(t) = V_\lambda(t) dW(t) + \Big(1 + \frac{1-\kappa}{2} V_\lambda(t) - 2\lambda V_\lambda(t)^2\Big) dt,$$

the generator of the diffusion process $V_\lambda(t)$ is $\mathcal{L}_\lambda = d/m_\lambda(dx) \cdot d/dS_\lambda(x)$ where

$$S_\lambda(x) = \int_1^x y^{\kappa-1} \exp\Big(\frac{2}{y} + 4\lambda y\Big) dy, \tag{3.1}$$

$$m_\lambda(dx) = 2x^{-\kappa-1} \exp\Big(-\frac{2}{x} - 4\lambda x\Big) dx. \tag{3.2}$$

We also put

$$\tilde{V}_\lambda(t) = V_\lambda(A_\lambda^{-1}(t)), \quad Y_\lambda(t) = S_\lambda(\tilde{V}_\lambda(t)),$$

where $A_\lambda^{-1}(t)$ is the inverse function of $A_\lambda(s) = \int_0^s V_\lambda(u) du$. The generators of $\tilde{V}_\lambda(t)$ and $Y_\lambda(t)$ are denoted by $\tilde{\mathcal{L}}_\lambda$ and $\mathcal{L}_\lambda^0$ respectively. $\mathcal{L}_\lambda^0$ is then given by $\mathcal{L}_\lambda^0 = d/m_\lambda^0(dx) \cdot d/dx$ with

$$m_\lambda^0(dx) = 2\theta_\lambda(x)^{-2\kappa+1} \exp\Big\{-\frac{4}{\theta_\lambda(x)} - 8\lambda\theta_\lambda(x)\Big\} dx, \tag{3.3}$$

where $\theta_\lambda(x)$ is the inverse function of $S_\lambda(\cdot)$. The path space representations of the diffusion processes with generators $\mathcal{L}_\lambda$, $\tilde{\mathcal{L}}_\lambda$ and $\mathcal{L}_\lambda^0$ are denoted by $\{\omega(t), t \geqq 0, P_\lambda^x\}$, $\{\omega(t), t \geqq 0, \tilde{P}_\lambda^x\}$ and $\{\omega(t), t \geqq 0, P_\lambda^{0,x}\}$, respectively. The expectations with respect to P_λ^x, $\tilde{P}_\lambda^x$ and $P_\lambda^{0,x}$ will be denoted by E_λ^x, $\tilde{E}_\lambda^x$ and $E_\lambda^{0,x}$, respectively. We begin by proving the following lemma.

LEMMA 1. *For any $x > 0$*

(3.4) $$\lim_{\lambda\downarrow 0}\lambda^{1-\kappa}M_\lambda(\lambda^{-\kappa}x)=M(x),$$

where $M_\lambda(x)$ *is the associated function of* $m_\lambda^0(dx)$ *and*

$$M(x)=2\gamma(\rho^{-1}(x)),$$

$$\gamma(x)=\int_0^x y^{-\kappa}e^{-4y}dy,$$

$\rho^{-1}(x)$ *is the inverse function of* $\rho(x)=\int_0^x y^{\kappa-1}e^{4y}dy.$

PROOF. By an easy computation we have for $x>\lambda$

$$S_\lambda(\lambda^{-1}x)=\lambda^{-\kappa}\rho_\lambda(x),$$

where

$$\rho_\lambda(x)=\int_\lambda^x y^{\kappa-1}\exp\Big(\frac{2\lambda}{y}+4y\Big)dy.$$

It is also easy to see that $\rho_\lambda(x)\to\rho(x)$ as $\lambda\downarrow 0$ and hence $S_\lambda(\lambda^{-1}x)\sim\lambda^{-\kappa}\rho(x)$ as $\lambda\downarrow 0$. Therefore for any $\varepsilon\in(0,x)$ $S_\lambda(\lambda^{-1}(x-\varepsilon))<\lambda^{-\kappa}\rho(x)$ holds for all sufficiently small $\lambda>0$. In other words

(3.5) $$\lambda^{-1}(x-\varepsilon)<\theta_\lambda(\lambda^{-\kappa}\rho(x))\ \text{for all sufficiently small}\ \lambda>0.$$

Similarly

(3.6) $$\lambda^{-1}(x+\varepsilon)>\theta_\lambda(\lambda^{-\kappa}\rho(x))\ \text{for all sufficiently small}\ \lambda>0.$$

Next we note that

$$M_\lambda(x)=2\int_1^{\theta_\lambda(x)}z^{-2\kappa+1}\exp\Big(-\frac{4}{z}-8\lambda z\Big)S_\lambda'(z)dz,$$

and hence

$$M_\lambda(\lambda^{-\kappa}\rho(x))=2\int_1^{\theta_\lambda(\lambda^{-\kappa}\rho(x))}z^{-\kappa}\exp\Big(-\frac{2}{z}-4\lambda z\Big)dz.$$

This combined with (3.5) and (3.6) yields

(3.7) $$2\int_1^{\lambda^{-1}(x-\varepsilon)}z^{-\kappa}\exp\Big(-\frac{2}{z}-4\lambda z\Big)dz<M_\lambda(\lambda^{-\kappa}\rho(x))<2\int_1^{\lambda^{-1}(x+\varepsilon)}z^{-\kappa}\exp\Big(-\frac{2}{z}-4\lambda z\Big)dz$$

for all sufficiently small $\lambda>0$. Since

$$2\int_1^{\lambda^{-1}x}z^{-\kappa}\exp\Big(-\frac{2}{z}-4\lambda z\Big)dz=2\lambda^{\kappa-1}\gamma_\lambda(x)$$

where

$$\gamma_\lambda(x)=\int_\lambda^x y^{-\kappa}\exp\Big(-\frac{2\lambda}{y}-4y\Big)dy,$$

(3.7) yields

$$\gamma_\lambda(x-\varepsilon) < M_\lambda(\lambda^{-\kappa}\rho(x))/(2\lambda^{\kappa-1}) < \gamma_\lambda(x+\varepsilon),$$

which again yields

$$\lim_{\lambda\downarrow 0} M_\lambda(\lambda^{-\kappa}\rho(x))/(2\lambda^{\kappa-1}) = \gamma(x),$$

because $\gamma_\lambda(x)\to\gamma(x)$ as $\lambda\downarrow 0$. Taking $\rho^{-1}(x)$ instead of x we obtain

$$\lim_{\lambda\downarrow 0} M_\lambda(\lambda^{-\kappa}x)/(2\lambda^{\kappa-1}) = \gamma(\rho^{-1}(x)),$$

which proves Lemma 1.

By virtue of Kasahara's continuity theorem and (2.7) Lemma 1 yields the following lemma concerning the characteristic functions $h_\lambda(\alpha)$ and $h(\alpha)$ of $M_\lambda(x)$ and $M(x)$ respectively.

LEMMA 2. $\lim_{\lambda\downarrow 0} \lambda^\kappa h_\lambda(\lambda\alpha)=h(\alpha),\ \alpha>0.$

Let $\tau=\inf\{t>0: \omega(t)=1\}$. Then for $x>0$ and $\alpha\geqq 0$

$$E_\lambda^x\left\{\exp\left(-\alpha\lambda\int_0^\tau \omega(s)ds\right)\right\} = \tilde{E}_\lambda^x\{e^{-\alpha\lambda\tau}\} = E_\lambda^{0,S_\lambda(x)}\{e^{-\alpha\lambda\sigma}\}, \tag{3.8}$$

where $\sigma=\inf\{t>0: \omega(t)=0\}$. For a given $a>0$ determine $a_\lambda(>1)$ by $S_\lambda(a_\lambda)=a$. Also let

$$S(x)=\int_1^x y^{\kappa-1}\exp\left(\frac{2}{y}\right)dy,$$

and determine $a_0(>1)$ by $S(a_0)=a$. Then $a_\lambda\uparrow a_0$ as $\lambda\downarrow 0$.

LEMMA 3. *If* $0<\kappa<1$, *then for* $\alpha>0$

$$1-E_\lambda^{a\lambda}\left\{\exp\left(-\alpha\lambda\int_0^\tau \omega(s)ds\right)\right\} \sim c(a,\alpha)\lambda^\kappa \quad \text{as } \lambda\downarrow 0, \tag{3.9}$$

where

$$c(a,\alpha)=a/h(\alpha). \tag{3.10}$$

PROOF. Since ∞ is not regular for $\mathcal{L}_\lambda^0$, (2.8) implies that for each $\alpha>0$ a positive decreasing solution $u_\lambda(\cdot,\alpha)$ of $\mathcal{L}_\lambda^0 u=\alpha u$ with $u(0)=1$ is unique and expressed as

$$u_\lambda(x,\alpha)=\varphi_\lambda(x,\alpha)-\psi_\lambda(x,\alpha)/h_\lambda(\alpha), \tag{3.11}$$

where the pair $\{\varphi_\lambda(x,\alpha),\psi_\lambda(x,\alpha)\}$ is the system of fundamental solutions associated with $m_\lambda^0(dx)$. By (3.8) we have

$$E_\lambda^{a\lambda}\left\{\exp\left(-\alpha\lambda\int_0^\tau \omega(s)ds\right)\right\} = u_\lambda(a,\alpha\lambda). \tag{3.12}$$

By making use of (2.3) and (2.4) with $m(dz)$ replaced by $m_\lambda^0(dz)$ we can easily prove that

$$\varphi_\lambda(a,\ \alpha\lambda)-1=2\alpha\lambda\int_0^a dx\int_1^{\theta_\lambda(x)} y^{-\kappa}\exp\left(-\frac{2}{y}-4\lambda y\right)dy+o(\lambda), \tag{3.13}$$

$$\psi_\lambda(a,\ \alpha\lambda)=a+O(\lambda), \tag{3.14}$$

as $\lambda\downarrow 0$. Denoting by $\theta(x)$ the inverse function of $S(\cdot)$, we have

$$\lim_{\lambda\downarrow 0}\frac{\varphi_\lambda(a,\ \alpha\lambda)-1}{\alpha\lambda}=2\int_0^a dx\int_1^{\theta(x)} y^{-\kappa}\exp\left(-\frac{2}{y}\right)dy=const.. \tag{3.15}$$

From (3.12) and (3.11) we have

$$1-E_\lambda^{a_\lambda}\left\{\exp\left(-\alpha\lambda\int_0^\tau \omega(s)ds\right)\right\}=1-\varphi_\lambda(a,\ \alpha\lambda)+\psi_\lambda(a,\ \alpha\lambda)/h_\lambda(\alpha\lambda),$$

which combined with (3.15), (3.14) and Lemma 2 finally implies (3.9).

The idea of the proof of (5) of Theorem 1 is as follows. We want to compute

$$\lim_{\lambda\downarrow 0}\mathcal{E}^0\{\exp(-\lambda T_{\lambda^{-\kappa}t})\}$$

which, by virtue of Kotani's formula, equals

$$\begin{aligned}(3.16)\qquad &\lim_{\lambda\downarrow 0}E\left\{\exp\left(-\int_0^{\lambda^{-\kappa}t}U_\lambda(s)ds\right)\right\}\\ &=\lim_{\lambda\downarrow 0}E\left\{\exp\left(-2\lambda\int_0^{\lambda^{-\kappa}t}V_\lambda(s)ds\right)\right\}\\ &=\lim_{\lambda\downarrow 0}E_\lambda^{\mu_\lambda}\left\{\exp\left(-2\lambda\int_0^{\lambda^{-\kappa}t}\omega(s)ds\right)\right\},\end{aligned}$$

where $E_\lambda^{\mu_\lambda}$ denotes the expectation with respect to $P_\lambda^{\mu_\lambda}=\int\mu_\lambda(dx)P_\lambda^x$, μ_λ being the invariant probability measure of the diffusion process with generator $\mathcal{L}_\lambda$. We first compute

$$\lim_{\lambda\downarrow 0}E_\lambda^{a_\lambda}\left\{\exp\left(-2\lambda\int_0^{\lambda^{-\kappa}t}\omega(s)ds\right)\right\},$$

with the starting point a_λ defined by $S_\lambda(a_\lambda)=a$, by showing that $\int_0^{\lambda^{-\kappa}t}\omega(s)ds$ can be approximated by $\sum\int_{\sigma_{k-1}}^{\tau_k}\omega(s)ds$ where σ_k and τ_k are defined as follows:

$$\sigma_0=0,\quad \tau_k=\inf\{t>\sigma_{k-1}:\omega(t)=1\},\quad k\geqq 1,$$

$$\sigma_k=\inf\{t>\tau_k:\omega(t)=a_\lambda\},\qquad k\geqq 1.$$

Note that $\sigma_k-\sigma_{k-1}$, $k\geqq 1$, are i.i.d. random variables.

LEMMA 4. (i) $m_\lambda = E^{a\lambda}_{\lambda|1}\{\sigma_1\} = a m_\lambda(\boldsymbol{R}^+) < \infty$, $\lambda \geqq 0$, *where* $\boldsymbol{R}^+ = (0, \infty)$.
(ii) *For any* $\varepsilon > 0$ *and* $\lambda \geqq 0$

$$P^{a\lambda}_\lambda\left\{\left|\frac{\sigma_n}{n} - m_\lambda\right| > \varepsilon\right\} \to 0 \quad \text{as } n \to \infty, \tag{3.17}$$

and the convergence is uniform in $\lambda \in [0, 1]$.

PROOF. For $b>0$ and $x>0$ let $I=(x, b)$ and $J=(0, y)$ or $I=(b, x)$ and $J=(y, \infty)$ according as $0<x<b$ or $0<b<x$. Let $\sigma = \inf\{t>0 : \omega(t)=b\}$. Then it is well-known that

$$E^x_\lambda\{\sigma\} = \int_I dS_\lambda(y) \int_J m_\lambda(dz) \tag{3.18}$$

holds (note that $-S_\lambda(0) = S_\lambda(\infty) = \infty$ is also taken into account in deriving the above formula). By virtue of (3.18) we can easily compute $m_\lambda = E^{a\lambda}_\lambda\{\tau_1\} + E^{a\lambda}_\lambda\{\sigma_1 - \tau_1\}$, obtaining (i). The assertion (3.17) is nothing but the law of large numbers for i.i.d. random variables. Only the uniform convergence needs proof for which it is enough to verify the uniform integrability of $\{\sigma_1, P^{a\lambda}_\lambda, 0 \leqq \lambda \leqq 1\}$, namely,

$$\lim_{N\to\infty} \sup_{0\leqq\lambda\leqq 1} \int_{\{\sigma_1 > N\}} \sigma_1 dP^{a\lambda}_\lambda = 0. \tag{3.19}$$

We use the fact that the diffusion process $\{\omega(t), t \geqq 0, P^{a\lambda}_\lambda\}$ can also be realized as a solution of the stochastic differential equation

$$dV(t) = V(t)dW(t) + \left\{1 + \frac{1-\kappa}{2} V(t) - 2\lambda V(t)^2\right\} dt \tag{3.20}$$

with $V(0) = a_\lambda$. Then a comparison theorem in stochastic differential equations implies that the solution of (3.20) lies below the solution of (3.20) with $\lambda = 0$. From this observation we see that

$$P^{a\lambda}_\lambda\{\tau_1 > N\} \leqq P^{a_0}_0\{\tau_1 > N\}, \tag{3.21}$$

and by a similar argument

$$P^{a\lambda}_\lambda\{\sigma_1 - \tau_1 > N\} \leqq P^1_1\{\sigma > N\}, \tag{3.22}$$

where $\sigma = \inf\{t>0 : \omega(t) = a_0\}$. Thus (3.19) follows from (3.21) and (3.22). The proof of the lemma is finished.

We are now in the final stage of the proof of Theorem 1 in the case $0<\kappa<1$. For fixed $t>0$ and small $\varepsilon>0$ we put

$$n_1(\lambda) = [\lambda^{-\kappa} t(1-\varepsilon)/m_0], \quad n_2(\lambda) = [\lambda^{-\kappa} t(1+\varepsilon)/m_0].$$

Then the uniform convergence of (3.17) implies

$$\lim_{\lambda\downarrow 0} P_\lambda^{a\lambda}\left\{\left|\frac{\sigma_{n_i(\lambda)}}{n_i(\lambda)}-m_\lambda\right|>m_0\varepsilon\right\}=0, \quad i=1, 2.$$

Since $m_\lambda \to m_0$ as $\lambda\downarrow 0$ we have

$$\lim_{\lambda\downarrow 0} P_\lambda^{a\lambda}\left\{\left|\frac{\sigma_{n_i(\lambda)}}{n_i(\lambda)}-m_0\right|>m_0\varepsilon\right\}=0, \quad i=1, 2, \tag{3.23}$$

and in particular

$$\lim_{\lambda\downarrow 0} P_\lambda^{a\lambda}\{\sigma_{n_1(\lambda)}<n_1(\lambda)m_0(1+\varepsilon)\}=1.$$

But $n_1(\lambda)m_0(1+\varepsilon)<\lambda^{-\kappa}t$ for all sufficiently small $\lambda>0$ and hence

$$\lim_{\lambda\downarrow 0} P_\lambda^{a\lambda}\{\sigma_{n_1(\lambda)}<\lambda^{-\kappa}t\}=1. \tag{3.24}$$

Similarly we have

$$\lim_{\lambda\downarrow 0} P_\lambda^{a\lambda}\{\sigma_{n_2(\lambda)}>\lambda^{-\kappa}t\}=1. \tag{3.25}$$

Next we put

$$Y_k=\int_{\sigma_{k-1}}^{\tau_k}\omega(s)ds, \quad Z_k=\int_{\tau_k}^{\sigma_k}\omega(s)ds, \quad k\geqq 1.$$

Then on the event $A_\lambda=\{\sigma_{n_1(\lambda)}<\lambda^{-\kappa}t<\sigma_{n_2(\lambda)}\}$

$$\sum_{k=1}^{n_1(\lambda)}(Y_k+Z_k)<\int_0^{\lambda^{-\kappa}t}\omega(s)ds<\sum_{k=1}^{n_2(\lambda)}(Y_k+Z_k)$$

holds and hence

$$\begin{aligned}&E_\lambda^{a\lambda}\left\{\exp\left(-2\lambda\sum_{k=1}^{n_2(\lambda)}(Y_k+Z_k)\right); A_\lambda\right\}\\ &\leqq E_\lambda^{a\lambda}\left\{\exp\left(-2\lambda\int_0^{\lambda^{-\kappa}t}\omega(s)ds\right); A_\lambda\right\}\\ &\leqq E_\lambda^{a\lambda}\left\{\exp\left(-2\lambda\sum_{k=1}^{n_1(\lambda)}(Y_k+Z_k)\right); A_\lambda\right\},\end{aligned}$$

where the notation $E_\lambda^{a\lambda}\{X; A\}$ stands for the integral of X over A with respect to $P_\lambda^{a\lambda}$. Since $P_\lambda^{a\lambda}\{A_\lambda\}\to 1$ by (3.24) and (3.25),

$$\begin{aligned}&E_\lambda^{a\lambda}\left\{\exp\left(-2\lambda\sum_{k=1}^{n_2(\lambda)}(Y_k+Z_k)\right)\right\}\\ &\leqq E_\lambda^{a\lambda}\left\{\exp\left(-2\lambda\int_0^{\lambda^{-\kappa}t}\omega(s)ds\right)\right\}+o(1)\\ &\leqq E_\lambda^{a\lambda}\left\{\exp\left(-2\lambda\sum_{k=1}^{n_1(\lambda)}(Y_k+Z_k)\right)\right\}+o(1),\end{aligned} \tag{3.26}$$

where $o(1)$ indicates a term which tends to 0 as $\lambda\downarrow 0$. We now compute

$$\lim_{\lambda\downarrow 0} E_{\lambda}^{a_\lambda}\Big\{\exp\Big(-2\lambda\sum_{k=1}^{n_i(\lambda)}(Y_k+Z_k)\Big)\Big\}$$
$$=\lim_{\lambda\downarrow 0} E_{\lambda}^{a_\lambda}\{\exp(-2\lambda Y_1)\}^{n_i(\lambda)}\cdot E_{\lambda}^{a_\lambda}\{\exp(-2\lambda Z_1)\}^{n_i(\lambda)}.$$

By Lemma 3 we have

$$\lim_{\lambda\downarrow 0}[E_{\lambda}^{a_\lambda}\{\exp(-2\lambda Y_1)\}]^{n_1(\lambda)}=\lim_{\lambda\downarrow 0}\{1-c(a,\,2)\lambda^{\kappa}\}^{n_1(\lambda)} \tag{3.27}$$
$$=\exp\{-c(a,\,2)t(1-\varepsilon)/m_0\}.$$

Similarly we have

$$\lim_{\lambda\downarrow 0} E_{\lambda}^{a_\lambda}\Big\{\exp\Big(-2\lambda\sum_{k=1}^{n_2(\lambda)}Y_k\Big)\Big\}=\exp\{-c(a,\,2)t(1+\varepsilon)/m_0\}. \tag{3.28}$$

On the other hand, we have

$$2\lambda\sum_{k=1}^{n_i(\lambda)}Z_k<2\lambda a_\lambda\sigma_{n_i(\lambda)}<2a_\lambda\lambda^{1-\kappa}t(1+\varepsilon)m_0^{-1}\sigma_{n_i(\lambda)}/n_i(\lambda)$$

which tends to 0 in probability as $\lambda\downarrow 0$ by virtue of $0<\kappa<1$ and (3.23). Therefore

$$\lim_{\lambda\downarrow 0} E_{\lambda}^{a_\lambda}\Big\{\exp\Big(-2\lambda\sum_{k=1}^{n_i(\lambda)}Z_k\Big)\Big\}=1. \tag{3.29}$$

From (3.26), (3.27), (3.28) and (3.29) we have for any $\varepsilon>0$

$$\exp\{-c(a,\,2)t(1+\varepsilon)/m_0\}\leqq\liminf_{\lambda\downarrow 0} E_{\lambda}^{a_\lambda}\Big\{\exp\Big(-2\lambda\int_0^{\lambda^{-\kappa}t}\omega(s)ds\Big)\Big\}$$
$$\leqq\limsup_{\lambda\downarrow 0} E_{\lambda}^{a_\lambda}\Big\{\exp\Big(-2\lambda\int_0^{\lambda^{-\kappa}t}\omega(s)ds\Big)\Big\}$$
$$\leqq\exp\{-c(a,\,2)t(1-\varepsilon)/m_0\},$$

which implies

$$\lim_{\lambda\downarrow 0} E_{\lambda}^{a_\lambda}\Big\{\exp\Big(-2\lambda\int_0^{\lambda^{-\kappa}t}\omega(s)ds\Big)\Big\}=\exp\{-c(a,\,2)t/m_0\}. \tag{3.30}$$

LEMMA 5. *For any* $t>0$

$$\lim_{\lambda\downarrow 0} E\Big\{\exp\Big(-\int_0^{\lambda^{-\kappa}t}U_\lambda(s)ds\Big)\Big\}=\exp\{-c(a,\,2)t/m_0\}. \tag{3.31}$$

PROOF. Note that by (3.16) the left hand side of (3.31) equals

$$\lim_{\lambda\downarrow 0} E_{\lambda}^{\mu_\lambda}\Big\{\exp\Big(-2\lambda\int_0^{\lambda^{-\kappa}t}\omega(s)ds\Big)\Big\}.$$

Let $\sigma_0=\inf\{t>0:\omega(t)=a_\lambda\}$ and put $\Gamma_u=\{\sigma_0<u\}$. Then using the strong Markov property we see that

$$E_\lambda^{\mu_\lambda}\left\{\exp\left(-2\lambda\int_0^{\lambda^{-\kappa}t}\omega(s)ds\right);\Gamma_u\right\}$$

is bounded from below by

$$E_\lambda^{\mu_\lambda}\left\{\exp\left(-2\lambda\int_0^{\sigma_0}\omega(s)ds\right);\Gamma_u\right\}\cdot E_\lambda^{a_\lambda}\left\{\exp\left(-2\lambda\int_0^{\lambda^{-\kappa}t}\omega(s)ds\right)\right\}$$

and is bounded from above by

$$E_\lambda^{a_\lambda}\left\{\exp\left(-2\lambda\int_0^{\lambda^{-\kappa}(t-\lambda^\kappa u)}\omega(s)ds\right)\right\}.$$

Taking into account of the fact that $P_\lambda^{\mu_\lambda}\{\Gamma_u\}\to 1$ as $u\to\infty$ uniformly in $\lambda\in(0,1)$ and also of (3.30), we first left $\lambda\downarrow 0$ and then $u\uparrow\infty$. As a result we obtain (3.31).

The proof of (5) in Theorem 1 is now completed as follows. By Kotani's formula we have for $\xi>0$

$$\mathcal{E}^0\{\exp(-\xi T_x/x^{1/\kappa})\} = E\left\{\exp\left(-\int_0^x U_{\xi x^{-1/\kappa}}(s)ds\right)\right\}$$
$$= E\left\{\exp\left(-\int_0^{\lambda^{-\kappa}t} U_\lambda(s)ds\right)\right\},$$

where we put $t=\xi^\kappa$ and $\lambda=\xi x^{-1/\kappa}$. Letting $x\to\infty$ (so $\lambda\downarrow 0$) we obtain

$$\lim_{x\to\infty}\mathcal{E}^0\{\exp(-\xi T_x/x^{1/\kappa})\} = \exp\{-c(a,2)t/m_0\} = e^{-ct} = e^{-c\xi^\kappa},$$

where $c=\{2^{1-\kappa}\Gamma(\kappa)h(2)\}^{-1}$, because $c(a,2)=a/h(2)$ and

$$m_0 = am_0(\boldsymbol{R}^+) = 2a\int_0^\infty x^{-\kappa-1}e^{-2/x}dx = 2^{1-\kappa}\Gamma(\kappa)a$$

by (3.10) and (i) of Lemma 4.

Finally we prove (6). Clearly we see that for any $u>0$, $y>0$ and $t>0$

$$\{T_u\geqq t\}\subset\{\bar{\omega}(t)\leqq u\}\subset\{\omega(t)\leqq u\}\subset\{\underline{\omega}(t)\leqq u\}$$
$$\subset\{T_{u+y}\geqq t\}\cup\left\{\inf_{s\geq T_{u+y}}\omega(s)-(u+y)\leqq -y\right\}.$$

We notice that

$$\lim_{y\to\infty}\mathcal{P}^0\left\{\inf_{s\geq T_{u+y}}\omega(s)-(u+y)\leqq -y\right\} = \lim_{y\to\infty}\mathcal{P}^0\left\{\inf_{s\geq 0}\omega(s)\leqq -y\right\} = 0 \tag{3.32}$$

since $\omega(s)\to\infty$ as $s\uparrow\infty$, $\mathcal{P}^0$-a.s.. Therefore we have for all $x>0$

$$\mathcal{P}^0\{T_{t^\kappa x}\geqq t\}\leqq\mathcal{P}^0\{\bar{\omega}(t)\leqq t^\kappa x\}\leqq\mathcal{P}^0\{\omega(t)\leqq t^\kappa x\} \tag{3.33}$$
$$\leqq\mathcal{P}^0\{\underline{\omega}(t)\leqq t^\kappa u\}\leqq\mathcal{P}^0\{T_{t^\kappa x+y}\geqq t\}+\mathcal{P}^0\left\{\inf_{s\geq 0}\omega(s)\leqq -y\right\}.$$

For any sufficiently large fixed y, the result (5) ensures

$$\lim_{t\to\infty}\mathcal{P}^0\{T_{t^\kappa x}\geqq t\} = \lim_{t\to\infty}\mathcal{P}^0\{T_{t^\kappa x+y}\geqq t\}$$
$$= \lim_{v\to\infty}\mathcal{P}^0\{v^{-1/\kappa}T_v\geqq x^{-1/\kappa}\} = 1-F_\kappa(x^{-1/\kappa}),$$

which combined with (3.32) and (3.33) proves (6).

§ 4. Proof of Theorem 2.

For the proof of (i) it is enough to show that

$$\text{(4.1)}\quad \lim_{\lambda\to\infty}\mathcal{P}^0\Big\{\exp\Big(-\sum_{k=1}^n \xi_k\lambda^{-1/\kappa}(T_{\lambda x_k}-T_{\lambda x_{k-1}})\Big)\Big\} = \exp\Big\{-c\sum_{k=1}^n (x_k-x_{k-1})\xi_k^\kappa\Big\}$$

for any $\xi_1, \xi_2, \cdots, \xi_n\geqq 0$ and $0=x_0<x_1<\cdots<x_n$. Take an ε such that $0<\varepsilon<\min\{x_k-x_{k-1}: 1\leqq k\leqq n\}$ and let us prove first that

$$\text{(4.2)}\quad \lim_{\lambda\to\infty}\mathcal{E}^0\Big\{\exp\Big(-\sum_{k=1}^n \xi_k\lambda^{-1/\kappa}(T_{\lambda(x_k-\varepsilon)}-T_{\lambda x_{k-1}})\Big)\Big\} = \exp\Big\{-c\sum_{k=1}^n (x_k-x_{k-1}-\varepsilon)\xi_k^\kappa\Big\}.$$

In what follows $T_x=T_x(\omega)$ denotes the first passage time $\inf\{t>0: \omega(t)=x\}$ where x may lie either to the right or to the left of $\omega(0)$. If we put

$$F_{\lambda,k} = E_W^{\lambda x_{k-1}}\{\exp(-\xi_k\lambda^{-1/\kappa}T_{\lambda(x_k-\varepsilon)});\ T_{\lambda(x_k-\varepsilon)}<T_{\lambda(x_{k-1}-\varepsilon)}\},$$
$$G_{\lambda,k} = E_W^{\lambda x_{k-1}}\{\exp(-\xi_k\lambda^{-1/\kappa}T_{\lambda(x_k-\varepsilon)});\ T_{\lambda(x_k-\varepsilon)}>T_{\lambda(x_{k-1}-\varepsilon)}\},$$

then

$$\text{(4.3)}\quad E_W^0\Big\{\exp\Big(-\sum_{k=1}^n \xi_k\lambda^{-1/\kappa}(T_{\lambda(x_k-\varepsilon)}-T_{\lambda x_{k-1}})\Big)\Big\}$$
$$= \prod_{k=1}^n E_W^{\lambda x_{k-1}}\{\exp(-\xi_k\lambda^{-1/\kappa}T_{\lambda(x_k-\varepsilon)})\}$$
$$= \prod_{k=1}^n (F_{\lambda,k}+G_{\lambda,k}).$$

Making use of a trivial inequality $0\leqq\prod_{k=1}^n(a_k+b_k)-\prod_{k=1}^n a_k\leqq\sum_{k=1}^n b_k$ which holds under the assumption that $a_k, b_k\geqq 0$ and $a_k+b_k\leqq 1$ $(1\leqq k\leqq n)$, we have

$$\text{(4.4)}\quad 0\leqq E\Big\{\prod_{k=1}^n (F_{\lambda,k}+G_{\lambda,k})\Big\}-E\Big\{\prod_{k=1}^n F_{\lambda,k}\Big\}$$
$$\leqq\sum_{k=1}^n E\{G_{\lambda,k}\}\leqq\sum_{k=1}^n \mathcal{P}^{\lambda x_{k-1}}\{T_{\lambda(x_k-\varepsilon)}>T_{\lambda(x_{k-1}-\varepsilon)}\}$$
$$=\sum_{k=1}^n E\Big\{\int_{\lambda x_{k-1}}^{\lambda(x_k-\varepsilon)} e^{W_\kappa(x)}dx\Big/\int_{\lambda(x_{k-1}-\varepsilon)}^{\lambda(x_k-\varepsilon)} e^{W_\kappa(x)}dx\Big\}$$

$$= \sum_{k=1}^{n} E\left\{\int_0^{\lambda(x_k-x_{k-1}-\varepsilon)} e^{W_\kappa(y+\lambda x_{k-1})}dy \Big/ \int_{-\lambda\varepsilon}^{\lambda(x_k-x_{k-1}-\varepsilon)} e^{W_\kappa(y+\lambda x_{k-1})}dy\right\}$$

$$= \sum_{k=1}^{n} E\left\{\int_0^{\lambda(x_k-x_{k-1}-\varepsilon)} e^{W_\kappa(y)}dy \Big/ \int_{-\lambda\varepsilon}^{\lambda(x_k-x_{k-1}-\varepsilon)} e^{W_\kappa(y)}dy\right\}$$

(since W_κ has stationary increments)

$$\leqq nE\left\{\int_0^{\infty} e^{W_\kappa(y)}dy \Big/ \int_{-\lambda\varepsilon}^{\infty} e^{W_\kappa(y)}dy\right\} \to 0, \quad \lambda\to\infty.$$

On the other hand it is easy to see that

$$E\left\{\prod_{k=1}^{n} F_{\lambda,k}\right\} = \prod_{k=1}^{n} E\{F_{\lambda,k}\}, \tag{4.5}$$

$$\begin{aligned}\lim_{\lambda\to\infty} E\{F_{\lambda,k}\} &= \lim_{\lambda\to\infty} E\{F_{\lambda,k}+G_{\lambda,k}\} \\ &= \lim_{\lambda\to\infty} \mathcal{E}^0\{\exp(-\xi_k\lambda^{-1/\kappa}T_{\lambda(x_k-x_{k-1}-\varepsilon)})\} \\ &= \exp\{-c(x_k-x_{k-1}-\varepsilon)\xi_k^\kappa\},\end{aligned} \tag{4.6}$$

the last equality being a consequence of (5). From (4.3)～(4.6) we obtain (4.2). To derive (4.1) from (4.2) it is enough to notice that

$$\begin{aligned} 0 &\leqq \mathcal{E}^0\left\{\exp\left(-\sum_{k=1}^{n}\xi_k\lambda^{-1/\kappa}(T_{\lambda(x_k-\varepsilon)}-T_{\lambda x_{k-1}})\right)\right\} \\ &\quad -\mathcal{E}^0\left\{\exp\left(-\sum_{k=1}^{n}\xi_k\lambda^{-1/\kappa}(T_{\lambda x_k}-T_{\lambda x_{k-1}})\right)\right\} \\ &\leqq \mathcal{E}^0\left\{1-\exp\left(-\sum_{k=1}^{n}\xi_k\lambda^{-1/\kappa}(T_{\lambda x_k}-T_{\lambda(x_k-\varepsilon)})\right)\right\} \\ &\leqq \sum_{k=1}^{n}\mathcal{E}^0\{1-\exp(-\xi_k\lambda^{-1/\kappa}(T_{\lambda x_k}-T_{\lambda(x_k-\varepsilon)}))\} \\ &= \sum_{k=1}^{n}\mathcal{E}^0\{1-\exp(-\xi_k\lambda^{-1/\kappa}T_{\lambda\varepsilon})\} \\ &\to \sum_{k=1}^{n}\{1-\exp(-c\varepsilon\xi_k^\kappa)\} \quad \text{as } \lambda\to\infty \\ &\to 0 \quad \text{as } \varepsilon\downarrow 0. \end{aligned}$$

The proof of (ii) can be done in a way similar to (6). In fact, as in (3.33) we have for any $t_k>0$, $x_k>0$ $(1\leqq k\leqq n)$ and $y>0$

$$\begin{aligned} &\mathcal{P}^0\{T_{\lambda^\kappa x_k}\geqq\lambda t_k,\ 1\leqq k\leqq n\} \\ &\leqq \mathcal{P}^0\{\lambda^{-\kappa}\omega(\lambda t_k)\leqq x_k,\ 1\leqq k\leqq n\} \\ &\leqq \mathcal{P}^0\{T_{\lambda^\kappa x_k+y}\geqq\lambda t_k,\ 1\leqq k\leqq n\}+n\mathcal{P}^0\left\{\inf_{s\geq 0}\omega(s)\leqq -y\right\}. \end{aligned}$$

Letting $\lambda\uparrow\infty$ in the above we obtain

$$\lim_{\lambda\to\infty}\mathcal{P}^0\{\lambda^{-\kappa}\omega(\lambda t_k)\leqq x_k,\ 1\leqq k\leqq n\}$$
$$=P\{L(x_k)\geqq t_k,\ 1\leqq k\leqq n\}=P\{L^{-1}(t_k)\leqq x_k,\ 1\leqq k\leqq n\}.$$

§5. Proof of Theorem 1 in the case $\kappa=1$.

Assume $\kappa=1$ and recall

$$E_W^0\{e^{-\lambda T_t}\}=\exp\left\{-2\lambda\int_0^t V_\lambda(s)ds\right\},\quad a.s., \tag{5.1}$$

where $V_\lambda(t)=(2\lambda)^{-1}U_\lambda(t)$ is a stationary diffusion process with generator

$$\mathcal{L}_\lambda=\frac{d}{m_\lambda(dx)}\cdot\frac{d}{dS_\lambda(x)},$$

wherein

$$S_\lambda(x)=\int_0^x\exp\Big(\frac{2}{y}+4\lambda y\Big)dy,\quad m_\lambda(dx)=2x^{-2}\exp\Big(-\frac{2}{x}-4\lambda x\Big)dx.$$

Once the following proposition is proved, (7) of Theorem 1 follows immediately from (5.1).

PROPOSITION 1. *For any* $\lambda>0$

$$(x\log x)^{-1}\int_0^x V_{\lambda(x\log x)^{-1}}(t)dt\to 2\quad in\ probability\ as\ x\to\infty. \tag{5.2}$$

Before proving this proposition we prepare three lemmas. We put for $\xi>0$

$$a_\xi=m_\xi(\boldsymbol{R}^+)^{-1}\int_0^\infty x\,m_\xi(dx),$$

$$u_\xi(x)=\int_0^x dS_\xi(y)\int_0^y(z-a_\xi)m_\xi(dz).$$

LEMMA 6. $a_\xi\sim-2\log\xi$ *as* $\xi\downarrow 0$.

PROOF. a_ξ can be expressed as $a_\xi=2m_\xi(\boldsymbol{R}^+)^{-1}(I_\xi+II_\xi)$ where

$$I_\xi=\int_0^N x^{-1}\exp\{-2x^{-1}-4\xi x\}dx,\quad II_\xi=\int_N^\infty x^{-1}\exp\{-2x^{-1}-4\xi x\}dx.$$

For fixed $N>0$ I_ξ remains bounded as $\xi\downarrow 0$ and $e^{-2/N}II'_\xi\leqq II_\xi\leqq II'_\xi$ where

$$II'_\xi=\int_N^\infty x^{-1}e^{-4\xi x}dx=\int_{4\xi N}^\infty y^{-1}e^{-y}dy\sim-\log\xi$$

as $\xi \downarrow 0$. Thus the lemma is proved since $m_\xi(\boldsymbol{R}^+)^{-1} \to 1$ as $\xi \downarrow 0$.

LEMMA 7. $\xi^2 \int_0^\infty u_\xi(x)^2 m_\xi(dx) \to 0$ *as* $\xi \downarrow 0$.

PROOF. In what follows *const.* means a constant which is independent of ξ but may vary from place to place. First we prove

$$0 < -u_\xi(x) \leqq const.\, x \log \frac{1}{\xi} \quad \text{for } 0 < x \leqq a_\xi. \tag{5.3}$$

The restriction $0 < x \leqq a_\xi$ implies

$$\begin{aligned} 0 < -u_\xi(x) &= 2\int_0^x \exp\{2y^{-1}+4\xi y\}\, dy \int_0^y (a_\xi - z) z^{-2} \exp\{-2z^{-1}-4\xi z\}\, dz \\ &\leqq 2a_\xi \int_0^x \exp\{2y^{-1}+4\xi y\}\, dy \int_0^y z^{-2} \exp\{-2z^{-1}-4\xi z\}\, dz. \end{aligned}$$

Since $0 < y < x < a_\xi \sim -2\log\xi$ as $\xi \downarrow 0$, we have $0 < \xi y < \xi a_\xi \to 0$ as $\xi \downarrow 0$. Therefore

$$-u_\xi(x) \leqq const.\, \log\frac{1}{\xi} \int_0^x \exp\{2y^{-1}\}\, dy \int_0^y z^{-2} \exp\{-2z^{-1}\}\, dz \leqq const.\, x \log \frac{1}{\xi}$$

as was to be proved. Next we prove

$$0 < -u_\xi(x) \leqq const.\, x \log\frac{1}{\xi} + const.\, \frac{1}{\xi} \log\frac{x}{a_\xi} \quad \text{for } x > a_\xi. \tag{5.4}$$

Since

$$-\int_0^y (z-a_\xi) m_\xi(dz) = \int_y^\infty (z-a_\xi) m_\xi(dz),$$

$-u_\xi(x)$ can be expressed as $-u_\xi(x) = 2(I+II)$, where

$$\begin{aligned} 0 < I &= 2\int_0^{a_\xi} \exp\{2y^{-1}+4\xi y\}\, dy \int_0^y (a_\xi - z) z^{-2} \exp\{-2z^{-1}-4\xi z\}\, dz \\ &\leqq const.\, x \log\frac{1}{\xi} \quad \text{(by (5.3))}, \end{aligned}$$

$$\begin{aligned} 0 < II &= 2\int_{a_\xi}^x \exp\{2y^{-1}+4\xi y\}\, dy \int_y^\infty (z - a_\xi) z^{-2} \exp\{-2z^{-1}-4\xi z\}\, dz \\ &\leqq 2\int_{a_\xi}^x \exp\{2y^{-1}+4\xi y\}\, dy \int_y^\infty z^{-1} \exp\{-4\xi z\}\, dz. \end{aligned}$$

If we put $g(y) = \int_y^\infty z^{-1} e^{-4\xi z} dz$, then

$$g(y)=\int_{4\xi y}^{\infty}u^{-1}e^{-u}du\leqq(4\xi y)^{-1}e^{-4\xi y},$$

and hence

$$II\leqq 2\int_{a_\xi}^{x}\exp\{2y^{-1}+4\xi y\}\cdot(4\xi y)^{-1}e^{-4\xi y}dy$$
$$\leqq const.\frac{1}{\xi}\int_{a_\xi}^{x}\frac{dy}{y}=const.\frac{1}{\xi}\log\frac{x}{a_\xi}.$$

This proves (5.4). We can now complete the proof of Lemma 7 as follows. From (5.3) and (5.4) we have

$$\xi^2\int_0^\infty u_\xi(x)^2m_\xi(dx)\leqq const.\Big(\xi\log\frac{1}{\xi}\Big)^2\int_0^{a_\xi}\exp\{-2x^{-1}-4\xi x\}dx$$
$$+const.\Big(\xi\log\frac{1}{\xi}\Big)^2\int_{a_\xi}^{\infty}\exp\{-2x^{-1}-4\xi x\}dx$$
$$+const.\int_{a_\xi}^{\infty}\Big(\log\frac{x}{a_\xi}\Big)^2x^{-2}\exp\{-2x^{-2}-4\xi x\}dx$$
$$\leqq const.\Big(\xi\log\frac{1}{\xi}\Big)^2\int_0^\infty e^{-4\xi x}dx$$
$$+const.\int_{a_\xi}^{\infty}\Big(\log\frac{x}{a_\xi}\Big)^2x^{-2}dx$$
$$=const.\,\xi\Big(\log\frac{1}{\xi}\Big)^2+\frac{const.}{a_\xi}\int_1^\infty(\log y)^2y^{-2}dy$$
$$\to 0\qquad\text{as }\xi\downarrow 0.$$

LEMMA 8. $\Big(\log\frac{1}{\xi}\Big)^{-1}\xi\int_0^\infty|u_\xi'(x)|^2x^2m_\xi(dx)\to 0$ as $\xi\downarrow 0$.

PROOF. First we prove

$$0<-u_\xi'(x)\leqq const.\log\frac{1}{\xi}\quad\text{for }0<x\leqq a_\xi.\tag{5.5}$$

Under the restriction $0<x\leqq a_\xi$ we have

$$0<-u_\xi'(x)=2\exp\{2x^{-1}+4\xi x\}\int_0^x(a_\xi-y)y^{-2}\exp\{-2y^{-1}-4\xi y\}dy$$
$$<2a_\xi\exp\{2x^{-1}+4\xi x\}\int_0^x y^{-2}\exp\{-2y^{-1}-4\xi y\}dy$$
$$\leqq const.\log\frac{1}{\xi}e^{2/x}\int_0^x y^{-2}e^{-2/y}dy=const.\log\frac{1}{\xi}.$$

Next we prove

(5.6) $$0 < -u'_\xi(x) \leqq const.\, e^{4\xi x}\varphi(4\xi x) \quad \text{for } x > a_\xi,$$

where

$$\varphi(x) = \int_x^\infty y^{-1}e^{-y}dy.$$

In fact, if $x > a_\xi$ then

$$\begin{aligned} 0 < -u'_\xi(x) &= 2\exp\{2x^{-1}+4\xi x\}\int_x^\infty (y-a_\xi)y^{-2}\exp\{-2y^{-1}-4\xi y\}\,dy \\ &\leqq const.\, e^{4\xi x}\int_x^\infty y^{-1}e^{-4\xi y}dy \\ &= const.\, e^{4\xi x}\varphi(4\xi x). \end{aligned}$$

Now the proof of Lemma 8 is completed as follows.

$$\begin{aligned} &\left(\log\frac{1}{\xi}\right)^{-1}\xi\int_0^{a_\xi} |u'_\xi(x)|^2 x^2 m_\xi(dx) \\ &\leqq const.\, \xi\log\frac{1}{\xi}\int_0^{a_\xi}\exp\{-2x^{-1}-4\xi x\}\,dx \quad \text{(by (5.5))} \\ &\leqq const.\, \xi\log\frac{1}{\xi}\cdot a_\xi \to 0 \quad \text{as } \xi \downarrow 0. \end{aligned}$$

To estimate the integral over (a_ξ, ∞) we notice that

$$\varphi(x) \sim \begin{cases} e^{-x}/x & \text{as } x \to \infty, \\ \log\dfrac{1}{x} & \text{as } x \downarrow 0, \end{cases}$$

and hence $e^x\varphi(x)^2 \in L^1(0, \infty)$. Therefore

$$\begin{aligned} &\left(\log\frac{1}{\xi}\right)^{-1}\xi\int_{a_\xi}^\infty |u'_\xi(x)|^2 x^2 m_\xi(dx) \\ &\leqq const.\, \xi\left(\log\frac{1}{\xi}\right)^{-1}\int_{a_\xi}^\infty e^{4\xi x}\varphi(4\xi x)^2 dx \quad \text{(by (5.6))} \\ &= const.\left(\log\frac{1}{\xi}\right)^{-1}\int_{4\xi a_\xi}^\infty e^x\varphi(x)^2 dx \\ &\to 0 \quad \text{as } \xi \downarrow 0. \end{aligned}$$

We now proceed to the proof of Proposition 1. Since $\mathcal{L}_\xi u_\xi = x - a_\xi$, an application of Itô's formula yields

$$u_\xi(V_\xi(t)) - u_\xi(V_\xi(0)) = \int_0^t u'_\xi(V_\xi(s))V_\xi(s)dW(s) + \int_0^t (V_\xi(s) - a_\xi)ds.$$

Putting $t=x$, $\xi=\lambda(x \log x)^{-1}$ and multiplying the both sides of the above by $(x \log x)^{-1}$ we have

$$(5.7) \qquad (x \log x)^{-1}\int_0^x (V_\xi(s)-a_\xi)ds$$
$$= (x \log x)^{-1}\{u_\xi(V_\xi(x))-u_\xi(V_\xi(0))\} -(x \log x)^{-1}\int_0^x u'_\xi(V_\xi(s))V_\xi(s)dW(s).$$

The distribution of $V_\xi(s)$ is $c_\xi m_\xi(dx)$ where c_ξ is the normalizing constant which tends to a finite value as $\xi \downarrow 0$. Therefore the second moment of the left hand side of (5.7) is dominated by

$$const.(x \log x)^{-2}\int_0^\infty u_\xi(y)^2 m_\xi(dy)+const.(x \log x)^{-2}x\int_0^\infty |u'_\xi(y)|^2 y^2 m_\xi(dy),$$

which tends to 0 as $x \to \infty$ by virtue of Lemma 7 and Lemma 8 because

$$(x \log x)^{-2}x \sim \frac{1}{\lambda}\left(\log\frac{1}{\xi}\right)^{-1}\xi \quad \text{as } x \to \infty.$$

This combined with

$$(x \log x)^{-1}\int_0^x a_\xi ds \to 2 \quad \text{as } x \to \infty$$

proves Proposition 1.

The assertion of (7) of Theorem 1 follows immediately from Proposition 1. The assertion (8) can be derived from (7) by a method similar to that used for deriving (6) in § 3.

§ 6. Proof of Theorem 1 in the case $\kappa>1$.

For any integer $k\geqq 1$ we put $\tau_k=T_k-T_{k-1}$. Then for any $\lambda_1, \lambda_2, \cdots, \lambda_n \geqq 0$ we have

$$(6.1) \qquad E_W^0\left\{\exp\left(-\sum_{k=1}^n \lambda_k\tau_k\right)\right\} = \prod_{k=1}^n E_W^{k-1}\{\exp(-\lambda_k\tau_k)\}$$
$$= \prod_{k=1}^n f(\Gamma_{k-1}W, \lambda_k),$$

where $f(W, \lambda)=E_W^0\{\exp(-\lambda T_1)\}$ and $\Gamma_x: \boldsymbol{W}\to\boldsymbol{W}$ (for fixed x) is defined by $(\Gamma_x W)(y)=W(x+y)-W(x)$ for any $y\in\boldsymbol{R}$. From (6.1) and the ergodicity of $\{\Gamma_x\}$ it follows that $\{\tau_k, k\geqq 1, \mathcal{P}^0\}$ is stationary and ergodic. Therefore

$$(6.2) \qquad \lim_{n\to\infty}\frac{T_n}{n} = \lim_{n\to\infty}\frac{\tau_1+\cdots+\tau_n}{n} = \mathcal{E}^0\{\tau_1\}, \quad \mathcal{P}^0\text{-a.s..}$$

The condition $\kappa>1$ is equivalent to the finiteness of $\mathcal{E}^0\{\tau_1\}$ as will be seen below. First we compute $E_W^0\{T_1\}$; the result is

$$E_W^0\{T_1\} = \int_0^1 dS_W(x)\int_{-\infty}^x m_W(dy).$$

Therefore

$$\begin{aligned}\mathcal{E}^0\{\tau_1\} &= E\{W_W^0(T_1)\}\\ &= 2\int_0^1 dx\int_{-\infty}^x E\{\exp\{W_\kappa(x)-W_\kappa(y)\}\}\,dy = \frac{4}{\kappa-1}.\end{aligned}$$

Thus (9) follows from (6.2). Next we prove (10). Clearly $\bar{\omega}(t)\to\infty$ as $t\uparrow\infty$ ($\mathcal{P}^0$-*a.s.*) and for any $\varepsilon>0$

$$\frac{T_{\bar{\omega}(t)}}{\bar{\omega}(t)} \leqq \frac{t}{\bar{\omega}(t)} < \frac{T_{\bar{\omega}(t)+\varepsilon}}{\bar{\omega}(t)}.$$

Letting $t\uparrow\infty$ in the above we obtain

$$\lim_{t\to\infty}\frac{t}{\bar{\omega}(t)} = \frac{4}{\kappa-1}, \quad \mathcal{P}^0\text{-}a.\text{s.},$$

which means that the left hand side of (10) equals $(\kappa-1)/4$, a.s.. To prove that the second and the third terms of (10) equal $(\kappa-1)/4$, a.s., we put $\theta(n)=(1-\varepsilon)\cdot(\kappa-1)n/4$ for an integer $n\geqq 1$ and for $0<\varepsilon<1$. Then

$$\begin{aligned}&\mathcal{P}^0\Big\{\inf_{s\geq T_{\theta(n)}}\omega(s)-\theta(n)<-\sqrt{n}\Big\}\\ &= E\Big\{P_W^{\theta(n)}\Big(\inf_{s\geq 0}\omega(s)-\theta(n)<-\sqrt{n}\Big)\Big\}\\ &= \mathcal{P}^0\Big\{\inf_{s\geq 0}\omega(s)<-\sqrt{n}\Big\}.\end{aligned}$$

The last term in the above is a general term of a convergent series by the result of [**9**]. Therefore an application of Borel-Cantelli's lemma yields

$$\inf_{s\geq T_{\theta(n)}}\omega(s)-\theta(n)\geqq -\sqrt{n} \tag{6.3}$$

for all sufficiently large n, $\mathcal{P}^0$-*a*.s.. Since $T_{\theta(n)}/n\to 1-\varepsilon$ as $n\to\infty$ ($\mathcal{P}^0$-*a*.s.), (6.3) implies

$$\inf_{s\geq n}\omega(s)-\theta(n)\geqq -\sqrt{n} \tag{6.4}$$

for all sufficiently large n, $\mathcal{P}^0$-*a*.s.. For $t>0$ we now take $n=n(t)$ such that $n\leqq t<n+1$. Then (6.4) implies

$$\theta(n)-\sqrt{n} \leqq \underline{\omega}(n) \leqq \underline{\omega}(t) \leqq \omega(t) \leqq \bar{\omega}(t) \tag{6.5}$$

for all sufficiently large t, $\mathcal{P}^0$-a.s.. Dividing (6.5) by t and then letting $t \uparrow \infty$ we finally see that the second and the third terms of (10) equal $(\kappa-1)/4$, $\mathcal{P}^0$-a.s..

§7. Remark to the case $\kappa=0$.

A limit theorem concerning $\bar{\omega}(t)=\max\{\omega(s): 0 \leqq s \leqq t\}$ was obtained by Kotani (1988), Ogura (1989) and Kawazu-Tanaka (1989) independently and by different methods (proofs were unpublished). The result is as follows. Let $t>0$ and put

$$W^{\#}(x) = W(x) - \min_{[x\wedge 0,\, x\vee 0]} W, \quad x \in \boldsymbol{R},$$

$$d_t^+ = \min\{x>0: W^{\#}(x)=t\}, \quad V_t^+ = \min_{[0,\, d_t^+]} W,$$

$$d_t^- = \max\{x<0: W^{\#}(x)=t\}, \quad V_t^- = \min_{[d_t^-,\, 0]} W,$$

and define b_t^+ and b_t^- by $W(b_t^\pm)=V_t^\pm$ (such $b_t^\pm$ are uniquely determined P-a.s. for each fixed $t>0$). Let

$$M_t^+ = \max_{[0,\, b_t^+]} W, \quad J_t^+ = M_t^+ \vee (V_t^+ + t),$$

$$M_t^- = \max_{[b_t^-,\, 0]} W, \quad J_t^- = M_t^- \vee (V_t^- + t),$$

and finally define $b_t^{\#}$ by

$$b_t^{\#} = \begin{cases} \min\{x>0: W^{\#}(x)=t\} & \text{if } J_t^+ < J_t^-, \\ \min\{x>0: W(x)=J_t^-\} & \text{if } J_t^+ > J_t^-. \end{cases}$$

Then the process $\{\lambda^{-2}\bar{\omega}(e^{\lambda t}), t>0, \mathcal{P}^0\}$ converges to $\{b_t^{\#}, t>0, P\}$ as $\lambda \to \infty$ in the sense of convergence of finite dimensional distributions. In particular, $(\log t)^{-2}\bar{\omega}(t)$ converges in law to $b_1^{\#}$ as $t \to \infty$ and

$$E\{e^{-\xi b_1^{\#}}\} = \int_0^1 E_+^x\{e^{-\xi T_1}\}\, dx, \quad \xi \geqq 0,$$

where E_+^x and T_1 denote the expectation and the first hitting time to 1, respectively, for the reflecting Brownian motion on $[0, \infty)$ starting at x.

The corresponding result in the case of random walks on $\{0, 1, 2, \cdots\}$ with reflecting barrier at 0 was obtained by Golosov [2].

References

[1] T. Brox, A one-dimensional diffusion process in a Wiener medium, Ann. Probab., 14 (1986), 1206-1218.

[2] A.O. Golosov, Limit distributions for random walks in random environments, Soviet Math. Dokl., **28** (1983), 18-22.

[3] A.O. Golosov, Localization of random walks in one-dimensional random environments, Comm. Math. Phys., **92** (1984), 491-506.

[4] K. Itô, Stochastic Processes II, Iwanami, 1958, (in Japanese).

[5] K. Itô and M. Nisio, On stationary solutions of a stochastic differential equation, J. Math. Kyoto Univ., **4**-1 (1964), 1-75.

[6] I.S. Kac and M.G. Krein, On the spectral functions of the string, Amer. Math. Soc. Transl. (2), **103** (1974), 19-102.

[7] Y. Kasahara, Spectral theory of generalized second order differential operators and its applications to Markov processes, Japan J. Math., **1** (1975), 67-84.

[8] K. Kawazu, Y. Tamura and H. Tanaka, Localization of diffusion processes in one-dimensional random environment, J. Math. Soc. Japan, **44** (1992), 515-550.

[9] K. Kawazu and H. Tanaka, On the maximum of a diffusion process in a drifted Brownian environment, In Séminaire de Probabilités, (eds. J. Azéma et al.), Lecture Notes in Math., **1557**, Springer-Verlag, 1993, 78-85.

[10] H. Kesten, M.V. Kozlov and F. Spitzer, A limit law for random walk in a random environment, Compositio Math., **30** (1975), 145-168.

[11] H. Kesten, The limit distribution of Sinai's random walk in random environment, Physica, **138A** (1986), 299-309.

[12] F.B. Knight, Characterization of the Lévy measure of inverse local times of gap diffusions, In Seminar on Stochastic Processes, Birkhäuser, 1981, 53-78.

[13] S. Kotani and S. Watanabe, Krein's spectral theory of strings and generalized diffusion processes, In Functional Analysis in Markov Processes, (ed. M. Fukushima), Lecture Notes in Math.. **923**, Springer-Verlag, 1982, 235-259.

[14] M.V. Kozlov, A random walk on the line with stochastic structure, Theory Probab. Appl., **18** (1973), 406-408.

[15] Y. Ogura, One-dimensional bi-generalized diffusion processes, J. Math. Soc. Japan, **41** (1989), 213-242.

[16] S. Schumacher, Diffusions with random coefficients, Particle Systems, Random Media and Large Deviations, Contemp. Math., **41** (1985), 351-356.

[17] Ya.G. Sinai, The limiting behavior of a one-dimensional random walk in a random medium, Theory Probab. Appl., **27** (1982), 256-268.

[18] F. Solomon, Random walks in a random environment, Ann. Probab., **3** (1975), 1-31.

[19] H. Tanaka, Localization of a diffusion process in a one-dimensional Brownian environment, Comm. Pure Appl. Math., **XLVII** (1994), 755-766.

[20] S. Watanabe, On time inversion of one-dimensional diffusion processes, Z. Wahrscheinlichtkeitstheorie verw. Gebiete, **31** (1975), 115-124.

Kiyoshi Kawazu
Department of Mathematics
Faculty of Education
Yamaguchi University
Yoshida, Yamaguchi 753
Japan

Hiroshi Tanaka
Department of Mathematics
Faculty of Science and Technology
Keio University
Hiyoshi, Yokohama 223
Japan

Pergamon

Chaos, Solitons & Fractals, Vol. 8, No. 11, pp. 1807–1816, 1997

0960-0779/97 $17.00 + 0.00

PII: S0960-0779(97)00029-5

Limit Theorems for a Brownian Motion with Drift in a White Noise Environment

HIROSHI TANAKA

Department of Mathematics, Faculty of Science and Technology, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama 233, Japan

Abstract—This paper discusses limit theorems for a diffusion analogue of Kesten–Kozlov–Spitzer's random walk in a random environment. The results obtained are similar to theirs but can be presented in a more explicit form by the use of Krein's spectral theory for one-dimensional generalized second-order differential operators of the form $d/dMd/dx$. © 1997 Elsevier Science Ltd

1. INTRODUCTION

The model of a random process in a random environment that we discuss in this paper has its origin on physical grounds in the works by Chernov [1] and Temkin [2]. Mathematical investigation of this model was then done by Kozlov [3], Kesten–Kozlov–Spitzer [4], Solomon [5] and Sinai [6], all in the framework of random walks. A diffusion analogue of Sinai's random walk was discussed by Brox [7] and Schumacher [8]. In this paper we discuss a diffusion model, namely, a model continuous in time and space (still in one dimension), which corresponds to what was discussed by Kesten–Kozlov–Spitzer in the framework of random walks. A rough description of our model is the following. Consider the heat or electric flow on a one-dimensional conductor such as a thin wire. Suppose the flow has a drift (or a bias) so that the movement takes place more likely to the right than to the left. In addition, the conductor contains some impurities located irregularly that cause fluctuation of the drift. We thus consider a particle performing Brownian motion with its own constant drift plus a random drift caused by impurities, or in a more mathematically idealized form, we consider the random process $X(t)$ described by the stochastic differential equation

$$dX(t) = dB(t) + \frac{\kappa}{4}dt - \tfrac{1}{2}w'(X(t))\,dt, \quad X(0) = 0, \tag{1}$$

where $B(t)$ is a standard one-dimensional Brownian motion, $\{w'(x), x \in \mathrm{R}\}$ is a white noise independent of $B(t)$ and κ is a non-negative constant. The problem is to know how the long-time behavior of $X(t)$ changes under the perturbation of the white noise term. The answer for this will vary with κ as in Kesten–Kozlov–Spitzer [4].

The case $\kappa = 0$ was treated by Brox [7] and Schumacher [8]; the limiting behavior of $X(t)$ as $t \to \infty$ is the same as that of Sinai's random walk, namely, the limit distribution of $(\log t)^{-2}X(t)$ as $t \to \infty$ exists. Kesten [9] proved that the limit distribution has a density

$$\frac{2}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}\exp\{-(2k+1)^2\pi^2|x|/8\}.$$

The purpose of this paper is to report detailed results in the case $\kappa > 0$. In this case one may anticipate the result for $X(t)$ from that of Kesten–Kozlov–Spitzer. However, to obtain exact and more elaborated results such as the determination of constants appearing in the limit, we need rigorous arguments. The result for $0 < \kappa < 1$ and part of the results for $\kappa > 2$ were obtained by Kawazu–Tanaka [10, 11]. For the case $\kappa = 0$, cf. Tanaka [12] and references therein. The proof of the results for $1 \le \kappa \le 2$ is new but only its outline will be given (see Section 3). Full proof will appear elsewhere. It is to be noted that, our problem being formulated in the framework of diffusion processes, we can make much use of methods in differential equations, in particular, Krein's spectral theory. Our results (see Section 1) naturally correspond to those of Kesten–Kozlov–Spitzer, but some new light is shed from the viewpoint of methods. In Section 4 we give an explicit (but complicated) representation of the constant $c(\kappa)$ appearing in the limit distributions.

2. MAIN RESULTS

We begin with giving a precise definition of $X(t)$ since the stochastic differential eqn (1) has no rigorous meaning because of the term $w'(X(t))\mathrm{d}t$. By a formal argument we see that a "symbolic solution $X(t)$ of eqn (1)" is a diffusion process with the generator

$$\mathscr{L}_w = \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{\kappa}{4}\frac{\mathrm{d}}{\mathrm{d}x} - \tfrac{1}{2}w'(x)\frac{\mathrm{d}}{\mathrm{d}x} = \tfrac{1}{2}\mathrm{e}^{w(x)-(\kappa x)/2}\frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{-w(x)+(\kappa x)/2}\frac{\mathrm{d}}{\mathrm{d}x}\right).$$

The latter expression of $\mathscr{L}_w$ has a rigorous meaning since it is of the form of Feller's canonical representation of a generator (for a quick view cf., for example, Section 2.10 of Nagasawa [13]). The process $X(t)$ is therefore defined as the diffusion process with generator $\mathscr{L}_w$ (defined by the latter expression)—a Markov process with transition probability $\exp(t\mathscr{L}_w)(x,\mathrm{d}y)$. To be more precise we introduce the space $W = \{w \in C(\mathrm{R})\colon w(0) = 0\}$ and the Wiener measure P on W. Then the formal derivative $w'(x)$, $x \in \mathrm{R}$, is a white noise defined on the probability space (W,P). For each $w \in W$ we denote by P_w the probability law of the diffusion process with generator $\mathscr{L}_w$ starting from 0. Therefore, if $\Omega = C[0,\infty)$ and if $w(t)$ denotes the value of $w(\in \Omega)$ at time $t \ge 0$, then $\{w(t), t \ge 0, P_w\}$ is a diffusion process with generator $\mathscr{L}_w$ starting from 0. Let $\mathscr{P}(\mathrm{d}w\mathrm{d}w) = P(\mathrm{d}w)P_w(\mathrm{d}w)$. We then regard the process $\{w(t), t \ge 0\}$ as defined on the probability space $(W \times \Omega, \mathscr{P})$. This is the process $X(t)$ in the introduction and is called a Brownian motion with drift in a white noise environment in this paper.

We are interested in the long-time behavior of $w(t)$ as $t \to \infty$, or more precisely speaking, the limit distributions of $\tau(x)$ and $w(t)$ under the suitable centering and scaling as $x, t \to \infty$, where

$$\tau(x) = \inf\{s > 0 : \omega(s) = x\}, \quad x \ge 0,$$

Our main results are the following in which $c(\kappa)$ denotes a suitable constant depending on κ.

(I) (Kawazu–Tanaka [10]) If $0 < \kappa < 1$, then

$$\lim_{x\to\infty}\mathscr{P}\left\{\frac{\tau(x)}{x^{1/\kappa}} \le t\right\} = F_\kappa(t), \quad t > 0,$$

$$\lim_{t\to\infty}\mathscr{P}\left\{\frac{\omega(t)}{t^\kappa} \le x\right\} = 1 - F_\kappa(x^{-1/\kappa}), \quad x > 0,$$

where

$$\int_0^\infty \mathrm{e}^{-\xi t}\,\mathrm{d}F_\kappa(t) = \exp\{-c(\kappa)\xi^\kappa\}, \quad \xi \ge 0.$$

(II) If $\kappa = 1$, then

$$\lim_{x\to\infty} \mathcal{P}\left\{\frac{\tau(x) - 4x\log x}{x} \le t\right\} = F_1(t), \quad t \in \mathbb{R},$$

$$\lim_{t\to\infty} \mathcal{P}\left\{\frac{\omega(t) - \bar{\mu}(t)}{t(\log t)^{-2}} \le x\right\} = 1 - F_1(-16x), \quad x \in \mathbb{R},$$

where

$$\bar{\mu}(t) = t(4\log t)^{-1}\left(1 + \frac{\log\log t + \log 4}{\log t}\right),$$

$$\int_{-\infty}^{\infty} e^{-\xi t}\, dF_1(t) = \exp\{c(1)\xi + 4\xi\log\xi\}, \quad \xi \ge 0.$$

(III) If $1 < \kappa < 2$, then

$$\lim_{x\to\infty} \mathcal{P}\left\{\frac{\tau(x) - mx}{x^{1/\kappa}} \le t\right\} = F_\kappa(t), \quad t \in \mathbb{R},$$

$$\lim_{t\to\infty} \mathcal{P}\left\{\frac{\omega(t) - m^{-1}t}{t^{1/\kappa}} \le x\right\} = 1 - F_\kappa(-m^{1+1/\kappa}x), \quad x \in \mathbb{R},$$

where $m = 4(\kappa - 1)^{-1}$ and

$$\int_{-\infty}^{\infty} e^{-\xi t}\, dF_\kappa(t) = \exp\{c(\kappa)\xi^\kappa\}, \quad \xi \ge 0.$$

(IV) If $\kappa = 2$, then

$$\lim_{x\to\infty} \mathcal{P}\left\{\frac{\tau(x) - 4x}{\sqrt{x\log x}} \le t\right\} = (2^8\pi)^{1/2} \int_{-\infty}^{t} e^{-2^{-8}s^2}\, ds, \quad t \in \mathbb{R},$$

$$\lim_{t\to\infty} \mathcal{P}\left\{\frac{\omega(t) - (t/4)}{\sqrt{t\log t}} \le x\right\} = (2\pi)^{-1/2} \int_{-\infty}^{x} e^{-y^2/2}\, dy, \quad x \in \mathbb{R}.$$

(V) Assume $\kappa > 2$.

(i) Environment-wise invariance principle [14].

(a) For almost all w (with respect to P) the process

$$\left\{\frac{\tau(\lambda x) - M(\lambda x)}{\sqrt{A\lambda}}, \; x \ge 0, \; P_w\right\}$$

converges in law to a Brownian motion as $\lambda \to \infty$ (in the sense of convergence of probability measures on the Skorohod space) where $A = 64(\kappa - 1)^{-2}(\kappa - 2)^{-1}$ and

$$M(x) = E_w\{\tau(x)\}$$
$$= 2\int_0^x dy \int_{-\infty}^{y} \exp\left\{w(y) - w(z) - \frac{\kappa(y - z)}{2}\right\} dz, \quad x \ge 0.$$

(In this theorem the definition of $\tau(x)$ is modified slightly so that $\tau(x)$ becomes right continuous in x, namely, $\tau(x) = \inf\{s > 0: \omega(s) > x\}$).

(b) For almost all w the process

$$\left\{\frac{\omega(\lambda t)-\mu(\lambda t)}{\sqrt{m^{-3}A\lambda}},\ t\geq 0,\ P_w\right\}$$

converges in law to a Brownian motion as $\lambda\to\infty$ where A is the same as in (a), $m=4(\kappa-1)^{-1}$ and $\mu(t)$ is the inverse function of $M(x)$.

(ii) Invariance principle in random environment [11]. Each of the processes

$$\left\{\frac{\tau(\lambda x)-\lambda mx}{\sqrt{C\lambda}}, x\geq 0, \mathscr{P}\right\},\ \left\{\frac{\omega(\lambda t)-\lambda m^{-1}t}{\sqrt{m^{-3}C\lambda}}, t>0, \mathscr{P}\right\}$$

converges in law to a Brownian motion as $\lambda\to\infty$ where m is the same as in (i) and $C=64\kappa(\kappa-1)^{-3}(\kappa-2)^{-1}>A$.

Let E denote the expectation with respect to $\mathscr{P}$. Then it is easy to see that $E\{\tau(x)\}<\infty$ iff $\kappa>1$ and $E\{\tau(x)^2\}<\infty$ iff $\kappa>2$ for $x>0$. This is a simple reason why the result is divided into the five cases (I)–(V).

3. KEY METHODS

Here are two key methods in our arguments. One is to use Kotani's formula which reduces our problem to the study of another diffusion process $X_\lambda(t)$ described by a certain stochastic differential equation with non-random coefficients. The other is to use Krein's spectral theory with which we can obtain some asymptotic properties of the Laplace transform of a certain hitting time of $X_\lambda(t)$.

3.1. Kotani's formula (unpublished; see Ref. [10] for a proof)

For $\lambda\geq 0$ and $t\geq 0$ (we use t instead of x)

$$E_w\{e^{-\lambda\tau(t)}\}=\exp\left\{-\int_0^t U_\lambda(s)\,ds\right\},\ P-\text{a.s.},\tag{2}$$

where $U_\lambda(t)$ is the unique stationary positive solution of

$$dU_\lambda(t)=U_\lambda(t)\,dw(t)+\left\{2\lambda+\frac{1-\kappa}{2}U_\lambda(t)-U_\lambda(t)^2\right\}dt.$$

For our later arguments it is convenient to consider $X_\lambda(t)=(2\lambda)^{-1}U_\lambda(t)$, $\lambda>0$. Kotani's formula eqn (2) then yields

$$E_w\{e^{-\lambda\tau(t)}\}=\exp\left\{-2\lambda\int_0^t X_\lambda(s)\,ds\right\},\ P-\text{a.s.},\tag{3}$$

and $X_\lambda(t)$ is a stationary diffusion process on $\mathbb{R}^+=(0,\infty)$ obtained as the unique stationary positive solution of

$$dX_\lambda(t)=X_\lambda(t)\,dw(t)+\left(1+\frac{1-\kappa}{2}X_\lambda(t)-2\lambda X_\lambda(t)^2\right)dt.\tag{4}$$

Note that eqns (3) and (4) hold for $\lambda\geq 0$. The generator $\mathscr{L}_\lambda$ of $X_\lambda(t)$ is given by

$$\mathscr{L}_\lambda=\frac{d}{m_\lambda(dx)}\frac{d}{ds_\lambda(x)},$$

where

$$s_\lambda(x) = \int_1^x y^{\kappa-1} \exp\left(\frac{2}{y} + 4\lambda y\right) \mathrm{d}y, \tag{5}$$

$$m_\lambda(\mathrm{d}x) = 2x^{-\kappa-1} \exp\left(-\frac{2}{x} - 4\lambda x\right) \mathrm{d}x, \ x > 0. \tag{6}$$

3.2. Krein's spectral theory (Ref. [15], see also Refs [16–18])

Here we sketch its main part. Let $\mathcal{M}$ be the class of functions $M:[0,\infty) \to [0,\infty]$ that are right continuous and non-decreasing. Note that $\mathcal{M}$ contains the function $\equiv \infty$. We put $M(0-) = 0$ and $l = l_M = \sup\{x: M(x) < \infty\}$. When $l > 0$ (i.e. $\mathcal{M}(x) \not\equiv \infty$) we consider

$$\varphi(x,\alpha) = 1 + \alpha \int_0^x \mathrm{d}y \int_{0-}^y \varphi(z,\alpha)\, \mathrm{d}M(z), \ 0 \le x < l,$$

$$h(\alpha) = \int_0^l \frac{\mathrm{d}x}{\varphi(x,\alpha)^2}, \ \alpha > 0, \ (< \infty \text{ if } M(x) \not\equiv 0).$$

When $l = 0$ (i.e. $M(x) \equiv \infty$) we put $h(\alpha) \equiv 0$. The function $h(\alpha)$ is called the characteristic function of $M(x)$ and the correspondence $M(x) \leftrightarrow h(\alpha)$ is called Krein's correspondence. When $M(x) \not\equiv 0$, $h(\alpha)$ can be represented as

$$h(\alpha) = c + \int_{0-}^\infty \frac{\sigma(\mathrm{d}\xi)}{\alpha + \xi}, \ \alpha > 0, \tag{7}$$

where $c \ge 0$ and $\sigma(\mathrm{d}\xi)$ is a measure on $[0,\infty)$ such that

$$0 < c + \int_{0-}^\infty \frac{\sigma(\mathrm{d}\xi)}{1 + \xi} < \infty.$$

Let $\mathcal{H}$ be the class of functions h of the form of eqn (7) augmented by the functions that are identically equal to 0 or to ∞. Then one of the main results in Krein's theory is that the correspondence

$$M \in \mathcal{M} \leftrightarrow h \in \mathcal{H}$$

is one-to-one, onto and induces a homeomorphism under suitable topologies on $\mathcal{M}$ and $\mathcal{H}$. We now list up some facts that will be used in our arguments.

(i) For positive constants a and

$$\mathrm{acM(ax)} \leftrightarrow \frac{1}{\mathrm{a}} \mathrm{h(c\alpha)}.$$

(ii) When $l = \infty$, $\lim_{\alpha \downarrow 0} 1/\{\alpha h(\alpha)\} = M(\infty -) = 1/\sigma(\{0\})$ (when $l < \infty$, the limit $= \infty$).
(iii) (Krein [19], see also [16]: p. 266, [18]: p. 239). Let $l = \infty$ and put

$$h^*(\alpha) = h(\alpha) - \frac{\sigma(\{0\})}{\alpha} = c + \int_{0+}^\infty \frac{\sigma(\mathrm{d}\xi)}{\alpha + \xi} \in \mathcal{H}.$$

Then

$$h^*(\alpha) \leftrightarrow M^*(x)$$

where

$$M^*(x) = M(\gamma^{-1}(x)) \Big/ \left(1 - \frac{M(\gamma^{-1}(x))}{M(\infty -)}\right),$$

$\gamma^{-1}(x)$ is the inverse function of

$$\gamma(t)=\int_0^t\left(1-\frac{M(s)}{M(\infty-)}\right)^2 \mathrm{d}s,\ l^*=\int_0^\infty\left(1-\frac{M(s)}{M(\infty-)}\right)^2 \mathrm{d}s.$$

4. OUTLINE OF PROOF

We give here an outline of the proof of (II) since (III) and (IV) can be proved in the same spirit as (II). The key points in the proof are to make use of Kotani's formula and fact (iii) in Section 2. From a technical viewpoint of our proof, it is convenient to regard x in the first limit theorem of (II) *as time* so we use habitual notation t instead of x.

From now on we assume $\kappa=1$ and put $\lambda=\xi t^{-1}$ where $\xi>0$ and $t>0$; ξ is fixed but later we let $t\to\infty$ so $\lambda\downarrow 0$. By Kotani's formula we have

$$E\left\{\exp\left(-\xi\frac{\tau(t)-4t\log t}{t}\right);A\right\}=E\{\mathrm{e}^{-V(t)};A\}\mathrm{e}^{4\xi\log t} \tag{8}$$

for any event A in the probability space (W,P), where

$$V(t)=2\lambda\int_0^t X_\lambda(s)\,\mathrm{d}s$$

and the notation $E\{X;A\}$ (resp. $E\{X;A\}$) stands for the integral of X over A with respect to $\mathscr{P}$ (resp. P). Given $a>0$ we define a_λ by $s_\lambda(a_\lambda)=a$ where $s_\lambda(x)$ is given by eqn (5) with $\kappa=1$. Obviously $a_\lambda>1$ and $a_\lambda\uparrow a_0$ as $\lambda\downarrow 0$. We define stopping times T_n, T_n', $n=0, 1,\ldots$, based on the process $X_\lambda(s)$ as follows:

$$T_0=\inf\{s>0{:}X_\lambda(s)=a_\lambda\},\ \ T_0'=\inf\{s>T_0{:}X_\lambda(s)=1\},$$

$$T_n=\inf\{s>T_{n-1}'{:}X_\lambda(s)=a_\lambda\},\ \ T_n'=\inf\{s>T_n{:}X_\lambda(s)=1\},\ \ n\ge 1.$$

Then T_n-T_{n-1}, $n=1, 2,\ldots$, are i.i.d. random variables with

$$E\{T_n-T_{n-1}\}=am_\lambda,$$

$$\mathrm{Var}\{T_n-T_{n-1}\}\le\text{const.},\ \ 0\le\lambda\le 1,$$

where m_λ is the total mass $m_\lambda(\mathbb{R}^+)$ of the measure defined by eqn (6) with $\kappa=1$ and const. is independent of λ. Therefore

$$P\left\{\left|\frac{T_n-T_0}{n}-am_\lambda\right|>\varepsilon\right\}\le\frac{\text{const.}}{\varepsilon^2 n},\ \ (\varepsilon>0). \tag{9}$$

Let $\varepsilon(t)=t^{-1/2}(\log t)^{1/4}$ and put

$$n_1(t)=\left[\frac{t(1-\varepsilon(t))}{am_\lambda}\right],\ \ n_2(t)=\left[\frac{t(1+\varepsilon(t))}{am_\lambda}\right].$$

Then $n_i(t)\sim t(am_0)^{-1}$ as $t\to\infty$ because $m_\lambda\to m_0$ as $\lambda\downarrow 0$. Therefore inequality (9) implies, for $i=1$ and 2,

$$P\left\{\left|\frac{T_{n_i(t)}-T_0}{n_i(t)}-am_\lambda\right|>\varepsilon(t)\right\}\to 0 \text{ as } t\to\infty. \tag{10}$$

We now put $A_t(\xi)=\{T_{n_1(t)}<t<T_{n_2(t)}\}$. Then, by making use of fact (10) we can prove that

$P\{A_t(\xi)\}\to 1$ as $t\to\infty$ for each fixed $\xi>0$. On the other hand, by the definition of $A_t(\xi)$, writing $T(n)$ instead of T_n we have

$$E\{e^{-V(T(n_2(t)))}; A(\xi)\}\, e^{4\xi\log t}\le E\{e^{-V(t)}; A_t(\xi)\}\, e^{4\xi\log t}$$
$$\le E\{e^{-V(T(n_1(t)))}; A_t(\xi)\}\, e^{4\xi\log t}. \quad (11)$$

Next we are going to prove, for $i=1$ and 2, that

$$\lim_{t\to\infty} E\{e^{-\alpha V(T(n_i(t)))}\}\, e^{4\alpha\xi\log t} = \exp\{2\alpha k(\alpha)\xi+4\alpha\xi\log\xi\},\ \alpha>0, \quad (12)$$

with some constant $k(\alpha)>0$. Making use of the strong Markov property of $X_\lambda(s)$ we see that

$$E\{e^{-\alpha V(T(n(t)))}\}=E_0(E_1E_2)^{n(t)},$$

where

$$E_0=E\Big\{\exp\Big(-2\lambda\alpha\int_0^{T_0}X_\lambda(s)\,ds\Big)\Big\},\ E_1=E\Big\{\exp\Big(-2\lambda\alpha\int_{T_0}^{T_0'}X_\lambda(s)\,ds\Big)\Big\},$$
$$E_2=E\Big\{\exp\Big(-2\lambda\alpha\int_{T_0'}^{T_1}X_\lambda(s)\,ds\Big)\Big\},$$

and the suffix i in $n_i(t)$ is suppressed to simplify the notation. Let $\tilde X_\lambda(s)=s_\lambda(X_\lambda(\Phi_\lambda^{-1}(s)))$ where $\Phi^{-1}{}_\lambda(s)$ is the inverse function of $\Phi_\lambda(\mathrm{r})=\int_0^{\mathrm{r}}X_\lambda(u)\,du$. Then $X_\lambda(s)$ is a stationary diffusion process in $\mathbb{R}$ with generator

$$\tilde{\mathscr{L}}_\lambda=\frac{d}{\mu_\lambda(dx)}\frac{d}{dx}$$

where

$$\mu_\lambda(dx)=2\theta_\lambda(x)^{-1}\exp\Big\{-\frac{4}{\theta_\lambda(x)}-8\lambda\theta_\lambda(x)\Big\}\,dx,$$
$$\theta_\lambda(x)=\text{the inverse function of } s_\lambda(\cdot).$$

Denote by $\Big\{\omega(t), t\ge 0, \tilde P_\lambda^x\Big\}$ the path-space representation of the diffusion process with generator $\tilde{\mathscr{L}}_\lambda$ starting from x. Then we can prove that

$$E_1=\tilde E_\lambda^a\{e^{-2\lambda\alpha\tilde\tau}\}=u_\lambda(a,2\lambda\alpha),$$

where $\tilde\tau=\inf\{s>0:\omega(s)=0\}$ and $u_\lambda(x,\alpha)\equiv\tilde E_\lambda^x\{e^{-\alpha\tilde\tau}\}$ is the unique positive decreasing solution of $\tilde{\mathscr{L}}_\lambda u=\alpha u$, $x\ge 0$ with $u(0)=1$. The solution $u_\lambda(x,\alpha)$ can be expressed as

$$u_\lambda(x,\alpha)=\varphi_\lambda(x,\alpha)-\frac{\psi_\lambda(x,\alpha)}{h_\lambda(\alpha)},\quad \alpha>0,$$

where $\varphi_\lambda(x,\alpha)$ and $\psi_\lambda(x,\alpha)$ are the solutions of

$$\varphi_\lambda(x,\alpha)=1+\alpha\int_0^x dy\int_0^y\varphi_\lambda(z,\alpha)\mu_\lambda(dz),$$

$$\psi_\lambda(x,\alpha)=x+\alpha\int_0^x dy\int_0^y\psi_\lambda(z,\alpha)\mu_\lambda(dz),$$

respectively, and $h_\lambda(\alpha)\leftrightarrow M_\lambda(x)\equiv\mu_\lambda([0,x])$. What we want to know is the behavior of

$1 - E_1$, namely, of $1 - u_\lambda(a,2\lambda\alpha)$ as $\lambda \downarrow 0$. For this it is enough to know the behavior of $h_\lambda(2\lambda\alpha)$ as $\lambda \downarrow 0$. By fact (iii) of Section 2 we see that

$$h_\lambda^*(\alpha) \equiv h_\lambda(\alpha) - \frac{1}{\alpha\mu_\lambda([0,\infty))} \leftrightarrow M_\lambda^*(x)$$

where the explicit form of $M_\lambda^*(x)$ is given in terms of $M_\lambda(x) \equiv \mu_\lambda([0,x])$. Now an essential step is the following Lemma 1, which is stated in general $(1 \le \kappa \le 2)$. We put

$$e_\kappa(x) = \int_x^\infty y^{-\kappa}\, e^{-4y}\, dy, \; f_\kappa(x) = \int_0^x y^{\kappa-1}\, e^{4y}\, dy, \tag{13}$$

$$g_\kappa(x) = 4\int_0^x \{e_\kappa(f_\kappa^{-1}(y))\}^2\, dy, \; f_\kappa^{-1} = \text{the inverse function}, \tag{14}$$

$$M_\lambda = \mu_\lambda([0,\infty)) = 2\int_1^\infty x^{-\kappa} \exp\left(-\frac{2}{x} - 4\lambda x\right) dx. \tag{15}$$

4.1. Lemma 1

(i) If $1 \le \kappa < 2$, then

$$\lim_{\lambda\downarrow 0} M_\lambda^{-2}\lambda^{\kappa-1} M_\lambda^*(M_\lambda^{-2}\lambda^{\kappa-2}x) = M^*(x), \; x > 0, \; x \ne l^*,$$

where

$$M^*(x) = \begin{cases} \dfrac{1}{2(e_\kappa \circ f_\kappa^{-1} \circ g_\kappa^{-1})(x)} & \text{for } 0 < x < l^*, \\ \infty & \text{for } x \ge l^*, \end{cases} \tag{16}$$

$$l^* = g_\kappa(\infty) < \infty.$$

Therefore

$$\lim_{\lambda\downarrow 0} M_\lambda^{-2}\lambda^{-(\kappa-2)} h_\lambda^*(\lambda\alpha) = h^*(\alpha), \; \alpha > 0, \tag{17}$$

where $h^*(\alpha) \leftrightarrow M^*(x)$.

(ii) If $\kappa = 2$, then

$$\lim_{\lambda\downarrow 0} \lambda \log\frac{1}{\lambda} M_\lambda^*\left(x \log\frac{1}{\lambda}\right) = M^*(x), \; x > 0, \; x \ne l^*,$$

where $l^* = 4/M_0^2$, and $M^*(x) = 0$ for $0 \le x < l^*$ and $= \infty$ for $x \ge l^*$. Therefore

$$\lim_{\lambda\downarrow 0} \left(\log\frac{1}{\lambda}\right)^{-1} h_\lambda^*(\lambda\alpha) = h^*(\alpha), \; \alpha > 0,$$

where $h^*(\alpha) \equiv l^* \leftrightarrow M^*(x)$.

Coming back to the case $\kappa = 1$, we now know the behavior of $h_\lambda^*(2\lambda\alpha)$ as $\lambda \downarrow 0$ from eqn (17) with $\kappa = 1$. Therefore we know the behavior of $1 - E_1$ as $\lambda \downarrow 0$ and similarly of $1 - E_2$. Thus we can obtain the following result:

$$E_1 E_2 = 1 - 4a\alpha\lambda \log\frac{1}{\lambda} + 2a\alpha k(\alpha)\lambda(1 + o(1)) \quad \text{as } \lambda \downarrow 0; \tag{18}$$

$$k(\alpha) = 2\int_0^1 (2 - e^{-4x} - e^{-2x} - e^{-4/x} - e^{-2/x})\frac{dx}{x} + 2\alpha h^*(2\alpha). \tag{19}$$

Noting that $E_0 \to 1$ as $t \to \infty$, we finally obtain

$$E\{e^{-\alpha V(T(n(t)))}\} e^{4\alpha\xi \log t} \sim (E_1 E_2)^{n(t)} e^{4\alpha\xi \log t}$$

$$\sim (1 + 2a\alpha k(\alpha)\lambda)^{n(t)} e^{4\alpha\xi \log t - 4a\alpha n(t)\lambda \log 1/\lambda}$$

$$\to \exp\{2\alpha k(\alpha)\xi + 4\alpha\xi \log \xi\} \quad \text{as } t \to \infty,$$

proving eqn (12). Combining eqn (12) with the fact that $P\{A_t(\xi)\} \to 1$ as $t \to \infty$, we have

$$\lim_{t\to\infty} E\{e^{-V(t)}; A_t(\xi)\} = \lim_{t\to\infty} E\{e^{-V(t)}; A_t(\xi) \cap A_t(\xi+1)\}$$

$$= \exp\{2k(1)\xi + 4\xi \log \xi\},$$

which again combined with eqn (8) and inequality (11) implies

$$\lim_{t\to\infty} E\left\{\exp\left(-\xi\frac{\tau(t) - 4t\log t}{t}\right); A_t(\xi)\right\} = \lim_{t\to\infty} E\left\{\exp\left(-\xi\frac{\tau(t) - 4t\log t}{t}\right); A_t(\xi) \cap A_t(\xi+1)\right\}$$

$$= \exp\{2k(1)\xi + 4\xi \log \xi\}.$$

This proves the first limit theorem of (II) with $c(\kappa) = 2k(1)$ in virtue of the following technical lemma.

4.2. Lemma 2

Suppose we are given a sequence of random variables $\{X_n, n = 1, 2, \ldots\}$ and a sequence of events $\{A_n(\xi), n = 1, 2, \ldots\}$ for each $\xi \geq 0$. We put

$$\varphi_n(\xi) = E\{e^{-\xi X_n}; A_n(\xi)\},$$

$$\psi_n(\xi) = E\{e^{-\xi X_n}; A_n(\xi) \cap A_n(\xi+1)\},$$

and assume that the following conditions are satisfied:

(i) $\lim_{n\to\infty} P\{A_n(\xi)\} = 1$ *for* $\xi \geq 0$.
(ii) There exists a continuous function $\varphi(\xi)$ on $[0,\infty)$ such that

$$\lim_{n\to\infty} \varphi_n(\xi) = \lim_{n\to\infty} \psi_n(\xi) = \varphi(\xi) \quad \text{for } \xi \geq 0.$$

Then X_n converges in law as $n \to \infty$ to some random variable X whose law is uniquely determined by $E\{e^{-\xi X}\} = \varphi(\xi)$, $\xi \geq 0$.

To close this section we remark that the second limit theorem of (II) can be obtained from the first one with a small amount of effort.

5. REMARKS

1. The limit distribution dF_κ is strictly stable if $\kappa \neq 1$ and stable if $\kappa = 1$. The limit distribution concerning $\omega(t)$ is also strictly stable except for the case $0 < \kappa \leq 1$. It is the Mittag–Leffler distribution if $0 < \kappa < 1$ and a stable distribution if $\kappa = 1$.
2. The constant $c(\kappa)$ appearing in dF_κ is given as follows.
   (I) If $0 < \kappa < 1$, then

$$c(\kappa) = \{2^{1-\kappa}\Gamma(\kappa)h(2)\}^{-1},$$

   where $h(\alpha)$ is the characteristic function of $M(x) = 2\gamma(f_\kappa^{-1}(x))$, $f_\kappa(x)$ is given by eqn (13) and $\gamma(x) = \int_0^x y^{-\kappa} e^{-4y}\, dy$ (see Ref. [10]).

(II) If $\kappa = 1$, then

$$c(1) = 4\int_0^1 (2 - e^{-4x} - e^{-2x} - e^{-4/x} - e^{-2/x})\frac{dx}{x} + 4h^*(2).$$

(III) If $1 < \kappa < 2$, then

$$c(\kappa) = 2\left\{\frac{8}{\kappa - 1}\int_0^\infty x^{-\kappa+1}e^{-4x}\,dx + 2h^*(2)\right\}/m_0$$

$$= \left\{\frac{2^{2\kappa}}{\kappa - 1}\Gamma(2-\kappa) + 4h^*(2)\right\}/\{2^{1-\kappa}\Gamma(\kappa)\}.$$

In (II) and (III) $h^*(\alpha)$ is the characteristic function of $M^*(x)$ defined by eqn (16).

REFERENCES

1. Chernov, A. A., Replication of a multicomponent chain by the lightning mechanism. *Biofizika*, 1967, **12**, 297–301.
2. Temkin, D. E., One-dimensional random walks in a two-component chain. *Dokl. Akad. Nauk SSSR*, 1972, **206**, 27–30.
3. Kozlov, M. V., Random walk on the line with stochastic structure. *Theor. Probab. Appl.*, 1973, **18**, 406–408.
4. Kesten, H., Kozlov, M. V. and Spitzer, F., A limit law for random walk in a random environment. *Composito Math.*, 1975, **30**, 145–168.
5. Solomon, F., Random walks in a random environment. *Ann. Probab.*, 1975, **3**, 1–31.
6. Sinai, Ya. G., The limiting behavior of a one-dimensional random walk in a random medium. *Theor. Probab. Appl.*, 1982, **27**, 256–268.
7. Brox, T., A one-dimensional diffusion process in a Wiener medium. *Ann. Probab.*, 1986, **14**, 1206–1218.
8. Schumacher, S., Diffusions with random coefficients. In *Contemporally Math, Particle Systems, Random Media and Large Deviations*, ed. R. Durrett. American Mathematical Society, New York, 1985, Vol. 41, pp. 351–356.
9. Kesten, H., The limit distribution of Sinai's random walk in random environment. *Physica*, 1986, **138A**, 299–309.
10. Kawazu, K., and Tanaka, H., A diffusion process in a Brownian environment with drift. *J. Math. Soc. Japan*, 1997, **49**, 189–211.
11. Kawazu, K. and Tanaka, H., Invariance principle for a Brownian motion with large drift in a white noise environment (preprint).
12. Tanaka, H., Diffusion processes in random environments. In *Proc. International Congress of Mathematicians*, Zürich, 1994, Switzerland. Birkhäuser, Basel, 1995, pp. 1047–1054.
13. Nagasawa, M., *Schrödinger Equations and Diffusion Theory*. Birkhäuser, Berlin, 1993.
14. Tanaka, H., Environment-wise central limit theorem for a diffusion in a Brownian environment with large drift. In *Itô's Stochastic Calculus and Probability Theory*, eds N. Ikeda *et al.* Springer, Berlin, 1996, pp. 373–384.
15. Krein, M. G., On a generalization of an investigation of Stieltjes. *Dokl. Akad. Nauk SSR*, 1952, **87**, 881–884.
16. Dym, H. and McKean, H. P., *Gaussian Processes, Function Theory, and the Inverse Spectral Problem*. Academic Press, New York, 1976.
17. Kasahara, Y., Spectral theory of generalized second order differential operators and its applications to Markov processes. *Japanese J. Math.*, 1975, **1**, 67–84.
18. Kotani, S. and Watanabe, S., Krein's spectral theory of strings and generalized diffusion processes. In *Functional Analysis in Markov Processes, Lecture Notes in Math*, ed. M. Fukushima. Springer, Berlin, 1982, Vol. 923, pp. 235–259.
19. Krein, M. G., On some cases of the effective determination of the density of a non-homogeneous string from its spectral function. *Dokl. Akad. Nauk SSSR*, 1953, **93**, 617–620.

Hiroshima Math. J.
28 (1998), 129–137

Invariance principle for a Brownian motion with large drift in a white noise environment

Kiyoshi Kawazu and Hiroshi Tanaka
(Received November 14, 1996)

Abstract. This paper discusses an invariance principle for a Brownian motion with drift coefficient $\kappa/4$ in a white noise environment under the assumption that κ is large. Our method clarifies the relation between the environment-wise invariance principle discussed in [7] and the present result (the invariance principle in random environment).

Introduction

Let W be the space of continuous functions on $\mathbf{R}$ vanishing at 0 that is equipped with the Wiener measure P. For an element $w \in W$ let us denote by w_κ the element of W defined by $w_\kappa(x) = w(x) - (\kappa x/2)$ where κ is a given positive constant. For $w \in W$, P_w denotes the probability measure on $\Omega = C[0, \infty)$ such that $\mathbf{X}_x = \{\omega(t), t \geq 0, P_w\}$ is a diffusion process with generator

$$\mathscr{L}_w = \frac{1}{2} e^{w_\kappa}(x) \frac{d}{dx}\left(e^{-w_\kappa}(x)\frac{d}{dx}\right)$$

starting at 0, where $\omega(t)$ is the value of a function $\omega(\in \Omega)$ at time t. We regard $\omega(t)$ as a process defined on the probability space $\{W \times \Omega, \mathscr{P}\}$ where $\mathscr{P}(dw\, d\omega) = P(dw)P_w(d\omega)$. Then symbolically

$$d\omega(t) = dB(t) + \frac{\kappa}{4}\, dt - \frac{1}{2} w'(\omega(t))\, dt,$$

where $B(t)$ is a standard Brownian motion independent of the white noise $\{w'(x)\}$. We call the process $\mathbf{X} = \{\omega(t), t \geq 0, \mathscr{P}\}$ a Brownian motion with drift in a white noise environment; in [2] [6] [7] it is called a diffusion process in a Brownian environment with drift. The present authors obtained some limit theorems for $\mathbf{X}$ in [2] (see [8] for further results; see also [6] for a brief survey on related problems), which are analogous to those of [3] and [5]; however, some problems remain open. The present paper is a continuation of [7] and

1991 *Mathematics Subject Classification.* 60J60
Key words and phrases. Invariance principle, Brownian motion, Random environment.

discusses the central limit theorem, or more precisely speaking, invariance principle *in random environment* in the case $\kappa > 2$.

We set

$$M_x(= M(x)) = 2\int_0^x dy \int_{-\infty}^y e^{w_\kappa(y)-w_\kappa(z)}\,dz, x \in \mathbf{R},$$

$$\mu(t) = \text{the inverse function of } \{M_x, x \in \mathbf{R}\}, t \in \mathbf{R},$$

$$T_x(= T(x)) = \inf\{t > 0 : \omega(t) > x\}, x \geq 0.$$

We use also the following notation:

$$\bar{\omega}(t) = \max\{\omega(s) : 0 \leq s \leq t\}, \underline{\omega}(t) = \inf\{\omega(s) : s \geq t\}, (\omega \in \Omega),$$

$$\gamma = (\kappa - 1)/2, m = 4/(\kappa - 1) = 2/\gamma,$$

$$A = 64(\kappa - 1)^{-2}(\kappa - 2)^{-1} = 16\gamma^{-2}(2\gamma - 1)^{-1},$$

$$B = 64(\kappa - 1)^{-3}(\kappa - 2)^{-1} = 8\gamma^{-3}(2\gamma - 1)^{-1},$$

$$C = A + B = 64\kappa(\kappa - 1)^{-3}(\kappa - 2)^{-1} = 8(2\gamma + 1)\gamma^{-3}(2\gamma - 1)^{-1}.$$

The following theorem was proved in [7].

THEOREM A (*Environment-wise invariance principle, see* [7]). *When* $\kappa > 2$, *we have the following*:

(i) *For almost all* $w \in W$ *with respect to* P, *the process*

$$\left\{\frac{T_{\lambda x} - M_{\lambda x}}{\sqrt{A\lambda}}, x \geq 0, P_w\right\}$$

converges in law to a Brownian motion as $\lambda \to \infty$ (*in the sense of convergence of probability measures on the Skorohod space*).

(ii) *For almost all* w, *the process*

$$\left\{\frac{\omega(\lambda t) - \mu(\lambda t)}{\sqrt{m^{-3}A\lambda}}, t \geq 0, P_w\right\}$$

converges in law to a Brownian motion as $\lambda \to \infty$ (*in the sense of convergence of probability measures on* $C[0, \infty)$). *The same is true when* $\omega(\lambda t)$ *is replaced by either of* $\bar{\omega}(\lambda t)$ *and* $\underline{\omega}(\lambda t)$.

Our main theorems are the following ($\kappa > 2$ is assumed throughout).

THEOREM 1. (i) *The process*

$$\left\{\frac{M_{\lambda x} - \lambda m x}{\sqrt{B\lambda}}, x \in \mathbf{R}, P\right\}$$

converges in law to a Brownian motion as $\lambda \to \infty$.

(ii) *The process*

$$\left\{\frac{\mu(\lambda t)-\lambda m^{-1}t}{\sqrt{m^{-3}B\lambda}}, t\in\mathbf{R}, P\right\}$$

converges in law to a Brownian motion as $\lambda\to\infty$.

THEOREM 2 (*Invariance principle in random environment*). (i) *The process*

$$\left\{\frac{T_{\lambda x}-\lambda mx}{\sqrt{C\lambda}}, x\geq 0, \mathscr{P}\right\}$$

converges in law to a Brownian motion as $\lambda\to\infty$.

(ii) *The process*

$$\left\{\frac{\omega(\lambda t)-\lambda m^{-1}t}{m^{-3}C\lambda}, t\geq 0, \mathscr{P}\right\}$$

converges in law to a Brownian motion as $\lambda\to\infty$. *The same is true when* $\omega(\lambda t)$ *is replaced by either of* $\bar{\omega}(\lambda t)$ *and* $\underline{\omega}(\lambda t)$.

As in [7] we introduce a one parameter family of measure preserving transformations θ_t, $t\in\mathbf{R}$, on (W,P) defined by $(\theta_t w)(x)=w(x+t)-w(t)$, $x\in\mathbf{R}$. Clearly $\theta_t\theta_s=\theta_{t+s}$ and $\{\theta_t\}$ is ergodic. Set

$$f_0(w)=\int_{-\infty}^{0} e^{-w_\kappa(t)}\,dt. \tag{0.1}$$

Then $\theta_t f_0\equiv f_0(\theta_t w)=\int_{-\infty}^{t} e^{w_\kappa(t)-w_\kappa(s)}\,ds$ and we have the following (see [7]):

$$E_w\{T_x\}=M_x=2\int_0^x \theta_y f_0\,dy; \tag{0.2}$$

the first equality of (0.2) holds for $x\geq 0$ and the second one holds for $x\in\mathbf{R}$.

$$\mathrm{Var}_w\{T_x\}=8\int_0^x \theta_y g\,dy \text{ for } x\geq 0 \quad \left(g(w)=\int_{-\infty}^{0} e^{-w_\kappa(t)}(\theta_t f_0)^2\,dt\right). \tag{0.3}$$

$$E\{f_0\}=\gamma^{-1}, E\{f_0^2\}=2\gamma^{-1}(2\gamma-1)^{-1}. \tag{0.4}$$

$$E\{\mathrm{Var}_w(T_x)\}=Ax \text{ for } x\geq 0. \tag{0.5}$$

$$\mathrm{Var}\{M_x\}=Bx+O(1), x\to\infty, (\mathrm{Var}=\text{variance}). \tag{0.6}$$

It was also observed in [7] that

$$d\theta_t f_0=\theta_t f_0\,dw(t)-(\gamma\theta_t f_0-1)\,dt, \tag{0.7}$$

so $\theta_t f_0$ is a stationary diffusion process obtained as the unique stationary positive solution of the stochastic differential equation (0.7). Therefore

$$\theta_t f_0 - f_0 = \int_0^t \theta_s f_0 \, dw(s) - \lambda \int_0^t \theta_s f \, ds, \tag{0.8}$$

where $f = f_0 - \gamma^{-1}$.

1. Proof of Theorem 1

For the proof of Theorem 1 we need some lemmas.

LEMMA 1 ([7]). $t^{-1/2} \max\{\theta_s f_0 : |s| \le t\} \to 0$ as $t \to \infty$.

LEMMA 2 ([7]). *For any positive constants* c_1,

$$M_{t+u} - M_t = mu(1 + o(1)) + o(\sqrt{\lambda}), \quad |t| \le c_1 \lambda, \quad u \in \mathbf{R},$$

where $o(1)$ *represents a general term that tends to* 0 *as* $\lambda \to \infty$ *uniformly in* (t, u) *such that* $|t| \le c_1 \lambda$ *and* $u \in \mathbf{R}$, *for almost all* w; $o(\sqrt{\lambda})$ *is a term that can be expressed as* $o(1)\sqrt{\lambda}$.

To prove (i) of Theorem 1 it is enough to consider $\int_0^x \theta_y f \, dy$ by virtue of (0.2). Making habitual use of t to indicate time we write

$$\frac{1}{\sqrt{\lambda}} \int_0^{\lambda t} \theta_s f \, ds = \frac{1}{\gamma \sqrt{\lambda}} \int_0^{\lambda t} \theta_s f_0 \, dw(s) - \frac{1}{\gamma \sqrt{\lambda}} (\theta_{\lambda t} f_0 - f_0). \tag{1.1}$$

By the ergodicity of $\{\theta_t\}$ and also by (0.4) we see that the quadratic variation of the stochastic integral term in (1.1) tends to $Bt/4$ as $\lambda \to \infty$, a.s., so the stochastic integral term itself converges in law to $\{(B/4)^{1/2} w(t), t \ge 0, P\}$ as $\lambda \to \infty$. The second term of the right hand side of (1.1) is negligible by Lemma 1. Therefore $X_\lambda^+ = \{\lambda^{-1/2} \int_0^{\lambda t} \theta_s f \, ds, t \ge 0\}$ converges in law to $\{(B/4)^{1/2} w(t), t \ge 0\}$, so does $X_\lambda^- = \{\lambda^{-1/2} \int_0^{-\lambda t} \theta_s f \, ds, t \ge 0\}$ because of the reversibility of the diffusion $\theta_t f$. Now the assertion (i) of Theorem 1 follows from the fact that X_λ^+ and X_λ^- are asymptotically independent as $\lambda \to \infty$.

To proceed let $\xi = \lambda m$ and put

$$\beta_\lambda(t) = (B\lambda)^{-1/2}(M_{\lambda t} - \lambda m t), \quad \tilde{\beta}_\lambda(t) = (m^{-3} B \xi)^{-1/2}(\mu(\xi t) - \xi m^{-1} t). \tag{1.2}$$

Then the assertion of (ii) of Theorem 1 follows immediately from the following Lemma.

LEMMA 3. *For any* $t_0 > 0$ *and* $\varepsilon > 0$

$$\lim_{\lambda \to \infty} P\left\{ \sup_{|t| \le t_0} |\beta_\lambda(t) + \tilde{\beta}_\lambda(t)| > \varepsilon \right\} = 0.$$

PROOF. From the second equality of (1.2) we have $\lambda mt = M(\lambda t + \tilde{\beta}_\lambda(t)m^{-1}\sqrt{B\lambda})$ and hence

$$(B\lambda)^{-1/2}\{M(\lambda t + \tilde{\beta}_\lambda(t)m^{-1}\sqrt{B\lambda}) - M(\lambda t)\} = -\beta_\lambda(t),$$

so an application of Lemma 2 yields $\tilde{\beta}_\lambda(t)(1 + o(1)) + o(1) = -\beta_\lambda(t)$, where $o(1)$ is a term tending to 0 uniformly on each finite t-interval as $\lambda \to \infty$, a.s. This implies the lemma.

The following observation will lead to another proof of (ii) of Theorem 1. If $v(t)$ denotes the inverse function of $\int_0^x \theta_y f_0\, dy$, then $v(t) = \mu(2t)$ and the derivative $v'(t)$, which equals to $1/\theta_{v(t)} f_0$, is a stationary diffusion process obtained from $\theta_t f_0$ by changing time and scale.

2. The proof of Theorem 2

We give the proof of the part (i). Taking an arbitrary positive sequence $\{\lambda_n, n = 1, 2, \ldots\}$ tending to ∞, we denote by $P^{(n)}$ the probability law of the process $\{\lambda_n^{-1/2}(T_{\lambda_n x} - \lambda_n mx), x \geq 0, \mathscr{P}\}$. Note that $P^{(n)}$ is a probability measure on the Skorohod space $D = D[0, \infty)$. For the proof of the part (i) it is enough to show that $P^{(n)}$ converges to the probability law of the process $\{\sqrt{C}w(x), x \geq 0, P\}$ as $n \to \infty$. We first prove that the sequence $\{P^{(n)}, n \geq 1\}$ is tight. If $Q_w^{(n)}$ denotes the probability law of the process $\{\lambda_n^{-1/2}(T_{\lambda_n x} - M_{\lambda_n x}),$ $x \geq 0, P_w\}$, then $Q_w^{(n)} \to Q_1(P\text{-a.s.})$ by Theorem A and hence $Q^{(n)} = \int Q_w^{(n)} P(dw)$ also converges to Q_1 as $n \to \infty$ where Q_1 is the probability law (on D) of the process $\{\sqrt{A}w(x), x \geq 0, P\}$. All the convergence here is to be understood as the convergence of probability measures on D. Therefore for any $\varepsilon > 0$ there exists a compact set $K_1 \subset D$ such that $Q^{(n)}(K_1^c) < \varepsilon^2$ for all $n \geq 1$. We then have $P\{L_n\} < \varepsilon$ where $L_n = \{w : Q_w^{(n)}(K_1^c) > \varepsilon\} = \{w : Q_w^{(n)}(K_1) \leq 1 - \varepsilon\}$. We also introduce, for each fixed w, an element $\varphi_n(w)$ of Ω which is defined to be the function $\lambda_n^{1/2}(M_{\lambda_n x} - \lambda_n mx)$ of x. Then $P \circ \varphi_n^{-1} \to Q_2$ as $n \to \infty$ by Theorem 1 where Q_2 is the probability law (on Ω) of the process $\{\sqrt{B}w(x),$ $x \geq 0, P\}$. We can thus find a compact set $K_2 \subset \Omega$ such that $P \circ \varphi_n^{-1}(K_2^c) < \varepsilon$ for all $n \geq 1$. We now put $K = \{w_1 + w_2 : w_1 \in K_1, w_2 \in K_2\}$. Then K is a compact subset of D. Since

$$\frac{T_{\lambda_n x} - \lambda_n mx}{\sqrt{\lambda_n}} = \frac{T_{\lambda_n x} - M_{\lambda_n x}}{\sqrt{\lambda_n}} + \frac{M_{\lambda_n x} - \lambda_n mx}{\sqrt{\lambda_n}},$$

we have

$$\begin{aligned} P^{(n)}(K) &= \iint 1_K(w_1 + \varphi_n(w))Q_w^{(n)}(dw_1)P(dw) \\ &\geq \iint 1_{K_1}(w_1)1_{K_2}(\varphi_n(w))Q_w^{(n)}(dw_1)P(dw) \\ &= \int_{\varphi_n^{-1}(K_2)} Q_w^{(n)}(K_1)P(dw) \\ &\geq \int_{\varphi_n^{-1}(K_2)\cap L_n^c} (1-\varepsilon)P(dw) \geq (1-\varepsilon)(1-2\varepsilon), \end{aligned}$$

which proves that $\{P^{(n)}, n \geq 1\}$ is tight. Therefore, for the proof of Theorem 2 it is enough to show that

$$\lim_{n\to\infty} \int f(w)P^{(n)}(dw) = \iint f(w_1 + w_2)Q_1(dw_1)Q_2(dw_2), \tag{2.1}$$

for any function f of the form

$$f(w) = \exp\left\{\sqrt{-1}\sum_{j=1}^{k} \alpha_j w(t_j)\right\},$$

where $\alpha_j \in \mathbf{R}$ and $t_j \geq 0$, $1 \leq j \leq k$. For such an f the left hand side of (2.1) equals

$$\begin{aligned} &\lim_{n\to\infty} \iint f(w_1 + \varphi_n(w))Q_w^{(n)}(dw_1)P(dw) \\ &= \lim_{n\to\infty} \int_W \left\{\int_D f(w_1)Q_w^{(n)}(dw_1)\right\} f(\varphi_n(w))P(dw) \\ &= \int f(w_1)Q_1(dw_1) \int f(w_2)Q_2(dw_2) \end{aligned}$$

which also equals the right hand side of (2.1). This completes the proof of (i) of Theorem 2.

The part (ii) of Theorem 2 can be proved in a way similar to the above by making use of Theorem A and Theorem 1.

3. Supplement to the proof of (i) of Theorem A

The proof of Theorem A was given in [7]; however, some details in the proof of the part (i) were omitted. It will be worth supplementing them.

The proof of Theorem A given in [7] proceeds as follows. Let $\tau_k =$

$T_k - T_{k-1}$, $\tilde{\tau}_k = \tau_k - E_w\{\tau_k\}$, $k \geq 1$. Then it was proved that, for almost all w, $\{\tilde{\tau}_k, k \geq 1, P_w\}$ is a sequence of independent random variables satisfying the Lindeberg condition. Therefore the central limit theorem holds for T_n with respect to P_w, for almost all w. Note that $E_w\{T_n\} = M_n$ and $\text{Var}_w\{T_n\} \sim An$ as $n \to \infty$ (P-a.s.). Now the rest of the proof, whose detail was omitted in [7], is given as follows.

Let $t_{nk} = \text{Var}_w\{\tau_k\}/\text{Var}_w\{T_n\}$, $\zeta_{nk} = (An)^{-1/2}\{T_k - M_k\}$, $1 \leq k \leq n$, and $t_{n0} = \zeta_{n0} = 0$. For each fixed w we construct a piece-wise linear function $\xi_n(x)$, $0 \leq x \leq 1$, with vertexes $(\sum_{j=0}^{k} t_{nj}, \zeta_{nk})$, $0 \leq k \leq n$. We regard $\{\xi_n(x), 0 \leq x \leq 1, P_w\}$ as a process with time parameter x. Then by Theorem 3.1 of [5], for almost all w, the process $\{\xi_n(x), 0 \leq x \leq 1, P_w\}$ converges in law to a Brownian motion as $n \to \infty$. We now modify $\xi_n(x)$ slightly, namely, we consider a piece-wise linear function $\eta_n(x)$ with vertexes $(k/n, \zeta_{nk})$, $0 \leq k \leq n$. Then $\eta_n(x)$ can be represented as $\eta_n(x) = \xi_n(\varphi_n(x))$ where $\varphi_n(x)$ is the piece-wise linear function with vertexes $\left(k/n, \sum_{j=0}^{k} t_{nj}\right)$, $0 \leq k \leq n$. On the other hand it is easy to see that, for each fixed x, $\varphi_n \to x$ as $n \to \infty$ for almost all w. This combined with the fact that φ is increasing implies that $\varphi_n \to x$ uniformly as $n \to \infty$ (P-a.s.). Therefore

(3.1) the process $\{\eta_n(x), 0 \leq x \leq 1, P_w\}$ converges in law to a Brownian motion as $n \to \infty$ for almost all w.

We finally prove that the process $\{(A\lambda)^{-1/2}(T_{\lambda x} - M_{\lambda x}), x \in [0,1], P_w\}$ converges in law to a Brownian motion as $\lambda \to \infty$ for almost all w; the time interval $[0,1]$ can be replaced by an arbitrary interval $[0, t_0]$ with a minor modification of the proof. Given $x \in (0,1]$ and an integer $n \geq 1$ we take the integer k such that $(k-1)/n < x \leq k/n$. Then $T_{nx} - M_{nx} > T_{k-1} - M_k > \sqrt{An}\eta_n(x) - \tau_k - m_k$ where $m_k = M_k - M_{k-1}$. Similarly $T_{nx} - M_{nx} < \sqrt{An}\eta_n(x) + \tau_k + m_k$ and hence

$$\sqrt{An}\eta_n(x) - (\tau_k + m_k) < T_{nx} - M_{nx} < \sqrt{An}\eta_n(x) + (\tau_k + m_k).$$

This implies that for $x \in [0,1]$

$$\sqrt{An}\eta_n(x) - (\hat{\tau}_n + \hat{m}_n) < T_{nx} - M_{nx} < \sqrt{An}\eta_n(x) + (\hat{\tau}_n + \hat{m}_n), \tag{3.2}$$

where $\hat{\tau}_n = \max\{\tau_k : 1 \leq k \leq n\}$ and $\hat{m}_n = \max\{m_k : 1 \leq k \leq n\}$. Next, given $\lambda > 0$ let $n = n(\lambda)$ be the integer such that $n - 1 < \lambda \leq n$. Then $T_{(n-1)x} - M_{\lambda x} < T_{\lambda x} - M_{\lambda x} \leq T_{nx} - M_{\lambda x}$, which combined with (3.2) implies

$$\sqrt{A(n-1)}\eta_{n-1}(x) - (\hat{\tau}_{n-1} + \hat{m}_{n-1}) - (M_{\lambda x} - M_{(n-1)x})$$
$$< T_{\lambda x} - M_{\lambda x} < \sqrt{An}\eta_n(x) + (\hat{\tau}_n + \hat{m}_n) + (M_{nx} - M_{\lambda x}).$$

Since $M_{\lambda x} - M_{(n-1)x}$ and $M_{nx} - M_{\lambda x}$ are dominated by $2\hat{m}_n$, we obtain

$$(3.3) \quad \left(\frac{n-1}{\lambda}\right)^{1/2} \eta_{n-1}(x) - (A\lambda)^{-1/2}(\hat{\tau}_n + 3\hat{m}_n) < (A\lambda)^{-1/2}(T_{\lambda x} - M_{\lambda x})$$
$$< \left(\frac{n}{\lambda}\right)^{1/2} \eta_n(x) + (A\lambda)^{-1/2}(\hat{\tau}_n + 3\hat{m}_n).$$

On the other hand we can prove that for almost all w

$$(3.4) \quad P_w\left\{\lim_{n\to\infty} \hat{\tau}_n/\sqrt{n} = 0\right\} = 1,$$

$$(3.5) \quad \lim_{n\to\infty} \hat{m}_n/\sqrt{n} = 0.$$

In fact, it is easy to see that $\{\tau_k, k \geq 1, \mathscr{P}\}$ is stationary and ergodic. Since τ_k^2 is integrable we have $n^{-1}\sum_1^n \tau_k^2 \to$ const. as $n \to \infty$ ($\mathscr{P}$-a.s.) and hence $n^{-1}\tau_n^2 \to 0$, namely, $n^{-1/2}\tau_n \to 0$ ($\mathscr{P}$-a.s.). This implies $n^{-1/2}\hat{\tau}_n \to 0$ ($\mathscr{P}$-a.s.) and hence (3.4). (3.5) can be proved in a similar manner. By virtue of (3.1), (3.4) and (3.5) the processes of the leftmost and rightmost hands of (3.3) converge in law to a Brownian motion as $\lambda \to \infty$. Therefore (3.3) implies the assertion for $(A\lambda)^{-1/2}(T_{\lambda x} - M_{\lambda x})$ that we wanted to prove.

References

[1] K. Itô and H. P. McKean, Diffusion Processes and Their Sample Paths, Springer-Verlag, New York, 1965.

[2] K. Kawazu and H. Tanaka, A diffusion process in a Brownian environment with drift, J. Math. Soc. Japan **49** (1997), 189–211.

[3] H. Kesten, M. V. Kozlov and F. Spitzer, A limit law for random walk in a random environment, Composito Math. **30** (1975), 145–168.

[4] Yu. V. Prohorov, Convergence of random processes and limit theorems in probability theory, Theor. Probab. Appl. **1** (1956), 157–214.

[5] F. Solomon, Random walks in a random environment, Ann. Probab. **3** (1975), 1–131.

[6] H. Tanaka, Diffusion processes in random environments, Proc. International Congress of Mathematicians, Zürich, 1994, 1047–1054, Birkhäuser Verlag, Basel, 1995.

[7] H. Tanaka, Environment-wise central limit theorem for a diffusion in a Brownian environment with large drift, Itô's Stochastic Calculus and Probability Theory, ed. by N. Ikeda et al., pp. 373–384, Springer, 1996.

[8] H. Tanaka, Limit theorems for a Brownian motion with drift in a white noise environment, Chaos, Solitons & Fractals, **8** (1997), 1807–1816.

K. Kawazu
Department of Mathematics
Faculty of Education
Yamaguchi University
Yosida, Yamaguchi 753-8513
Japan

H. Tanaka
Department of Mathematics
Faculty of Science and Technology
Keio University
Yokohama 223-0061
Japan

Tanaka, H.
Osaka J. Math.
38 (2001), 369–377

SOME THEOREMS CONCERNING EXTREMA OF BROWNIAN MOTION WITH d-DIMENSIONAL TIME

Dedicated to Professor N. Ikeda on his 70th birthday

HIROSHI TANAKA

(Received October 16, 1999)

Introduction

Let $X = \{X(x), x \in \mathbf{R}^d\}$ be a Lévy's Brownian motion with d-dimensional time ([2]) defined on a certain probability space (Ω, P); thus X is a centered Gaussian system with continuous sample functions satisfying $X(0) = 0$ and $E\{X(x)X(y)\} = (|x| + |y| - |x - y|)/2$. For a nonempty subset A of $\mathbf{R}^d$ we put

$$\underline{X}(A) = \inf\{X(x) : x \in A\}, \quad \overline{X}(A) = \sup\{X(x) : x \in A\}.$$

We often use the notation $X(A)$ to denote either $\underline{X}(A)$ or $\overline{X}(A)$. For example, $X(A) - X(B)$ denotes any one of $\underline{X}(A) - \underline{X}(B)$, $\underline{X}(A) - \overline{X}(B)$, $\overline{X}(A) - \underline{X}(B)$ and $\overline{X}(A) - \overline{X}(B)$. A point x in $\mathbf{R}^d$ is called a point of local minimum (resp. local maximum) of a sample function X if there exists a neighborhood U of x such that $X(x) = \underline{X}(U)$ (resp. $X(x) = \overline{X}(U)$). A point of either local minimum or local maximum is called an extreme-point.

The following are typical of those problems and theorems we discuss in this paper.

(I) Under what condition on A does the probability distribution of $X(A)$ admit a strictly positive C^∞-density?

(II) Under what condition on A and B does the joint probability distribution of $X(A)$ and $X(B)$ admit a strictly positive C^∞-density?

(III) Almost all sample functions X have the following property: There are no distinct extreme-points x and y with $X(x) = X(y)$.

We give some sufficient conditions that will give positive answers to the problems (I) and (II) and then give a proof of (III). Formulating the problems somewhat generally we state our main results in the following theorems.

Theorem 1. *Let A_k, $1 \le k \le n$, be nonempty bounded closed sets not containing the origin 0. Then for any constants c_k, $1 \le k \le n$, such that $c_1 + c_2 + \cdots + c_n \ne 0$, the probability distribution of*

$$c_1 X(A_1) + c_2 X(A_2) + \cdots + c_n X(A_n)$$

can be expressed as a convolution $\gamma * \mu$ where γ is a nondegenerate Gaussian distribution with mean 0 and μ is some probability distribution in $\boldsymbol{R}$. In particular, the distribution of each of $\underline{X}(A)$ and $\overline{X}(A)$ has a strictly positive C^∞-density provided that A is a nonempty bounded closed set not containing 0.

Theorem 2. *Let A_j, B_k, $1 \le j \le m$, $1 \le k \le n$, be nonempty bounded closed sets such that $\cup_{j=1}^m A_j$ is separated from $\cup_{k=1}^n B_k$ by a certain $(d-1)$-dimensional hyperplane Π passing through the origin 0. Then for any constants c_j, c'_k, $1 \le j \le m$, $1 \le k \le n$, such that $\sum_{j=1}^m c_j \ne 0$ and $\sum_{k=1}^n c'_k \ne 0$, the joint distribution of*

$$(1) \qquad f_1(X) = \sum_{j=1}^{m} c_j X(A_j), \qquad f_2(X) = \sum_{k=1}^{n} c'_k X(B_k)$$

has a form $(\gamma_1 \otimes \gamma_2) * \nu$ where each γ_i is a nondegenerate Gaussian distribution with mean 0 and ν is some 2-dimensional probability distribution. In particular, the joint distribution of $X(A)$ and $X(B)$ has a strictly positive C^∞-density provided that A and B are nonempty bounded closed sets separated from each other by a certain $(d-1)$-dimensional hyperplane passing through 0.

Theorem 3. *Let A_j, B_k, $1 \le j \le m$, $1 \le k \le n$, be nonempty bounded closed sets such that $\cup_{j=1}^m A_j$ is separated from $\cup_{k=1}^n B_k$ by a certain $(d-1)$-dimensional hyperplane. Then for any constants c_j, c'_k, $1 \le j \le m$, $1 \le k \le n$, such that $\sum_{j=1}^m c_j = \sum_{k=1}^n c'_k \ne 0$, the probability distribution of $f_1(X) - f_2(X)$, with f_1 and f_2 given by (1), has a form $\gamma * \mu$ where γ is a nondegenerate Gaussian distribution with mean 0 and μ is some distribution in R. In particular, the distribution of each of $\underline{X}(A) - \underline{X}(B)$, $\underline{X}(A) - \overline{X}(B)$ and $\overline{X}(A) - \overline{X}(B)$ has a strictly positive C^∞-density provided that A and B are nonempty bounded closed sets separated from each other by a certain $(d-1)$-dimensional hyperplane.*

Theorem 4. *Almost all sample functions X have the following property: There. are no distinct extreme-points x and y of X such that $X(x) = X(y)$.*

An example of the applicability (or our motivation) of Theorem 4 will be given in the final section.

1. A lemma

Given a centered Gaussian system $\{X_\lambda,\ \lambda \in \Lambda\}$ defined on a certain probability space (Ω, P), we denote by H the real Hilbert space spanned by $\{X_\lambda,\ \lambda \in \Lambda\}$ and by H_0 the closed linear span (abbreviation: c.l.s.) of $\{X_\lambda - X_\mu,\ \lambda,\ \mu \in \Lambda\}$. Clearly $H_0 \subset H \subset L^2(\Omega, P)$. We now introduce the following conditions.

Condition (A). There exists a nondegenerate Gaussian random variable Y_0 independent of $\{X_\lambda - Y_0,\ \lambda \in \Lambda\}$.

Condition (B). There exists $\lambda \in \Lambda$ such that $X_\lambda \notin H_0$.

It is easy to see that the condition (B) implies that $X_\lambda \notin H_0$ for all $\lambda \in \Lambda$. Denote by $\mathbf{R}^\Lambda$ the space of real valued functions on Λ; it has a Borel structure defined in a natural way. Then we can regard $X_\Lambda = \{X_\lambda,\ \lambda \in \Lambda\}$ as a random variable taking values in $\mathbf{R}^\Lambda$. The following lemma is rather trivial; nevertheless, it plays a fundamental role in this paper.

Lemma 1. (i) *Let f be a Borel function from $\mathbf{R}^\Lambda$ to* R *such that*

$$f(w + t\mathbf{1}) = f(w) + ct \tag{1.1}$$

for any $w \in \mathbf{R}^\Lambda$ and $t \in$ R where c is some nonzero constant and $\mathbf{1}$ denotes the function on Λ that identically equals 1. *Then under the condition* (A) *we have $f(X_\Lambda) = cY_0 + Y$ with a suitable random variable Y independent of Y_0; in particular, the probability distribution of $f(X_\Lambda)$ has a strictly positive C^∞-density.*

(ii) *Suppose Λ is a locally compact space with a countable open base and assume that X_λ is continuous in λ with probability* 1. *We regard $X_\Lambda = \{X_\lambda,\ \lambda \in \Lambda\}$ as a random variable taking values in the space $C(\Lambda)$ of continuous functions on Λ, which is equipped with the compact uniform topology. Then, under the condition* (A), *the conclusion of* (i) *remains valid for any Borel function f from $C(\Lambda)$ to* R *satisfying* (1.1) *for $w \in C(\Lambda)$ and $t \in$* R.

(iii) *The condition* (B) *implies the condition* (A).

REMARK 1. Let Λ_k, $1 \le k \le n$, be subsets of Λ and let c_k, $1 \le k \le n$, be constants such that $c_1 + \cdots + c_n \ne 0$. Let $w(\Lambda_k)$ indicate either $\inf\{w(\lambda) : \lambda \in \Lambda_k\}$ or $\sup\{w(\lambda) : \lambda \in \Lambda_k\}$; the choice may depend on k but not on w. Then

$$f(w) = c_1 w(\Lambda_1) + \cdots + c_n w(\Lambda_n) \tag{1.2}$$

is a typical example of f satisfying (1.1) with $c = c_1 + \cdots + c_n$ provided that f can be defined to be a Borel function.

REMARK 2. Let F be a class of functions defined on [0, 1] and taking values in Λ (an example of such an F is the space of continuous paths in Λ connecting two given points of Λ). Then the function f defined by $f(w) = \inf\{g(w, u) : u \in F\}$ with $g(w, u) = \sup\{w(u(t)) : 0 \le t \le 1\}$ satisfies (1.1).

REMARK 3. If $\{X_\lambda,\ \lambda \in \Lambda\}$ satisfies (A) (resp. (B)) and if Λ_1 is a nonempty subset of Λ, then the sub-system $\{X_\lambda,\ \lambda \in \Lambda_1\}$ also satisfies (A) (resp. (B)).

Proof of Lemma 1. (i) Under the condition (A) $X_\Lambda - Y_0\mathbf{1}$ and Y_0 are independent so $f(X_\Lambda) - cY_0 = f(X_\Lambda - Y_0\mathbf{1})$ and Y_0 are independent. If we put $Y = f(X_\Lambda) - cY_0$,

then we have the expression $f(X_\Lambda) = cY_0 + Y$ in which Y_0 and Y are independent and Y_0 is a nondegenerate Gaussian random variable. The assertion (ii) follows from (i).
(iii) It is easy to see that $X_\lambda + H_0 = \{X_\lambda + Y : Y \in H_0\}$ does not depend on λ. The condition (B) means that $X_\lambda + H_0 \not\ni 0$. Since $X_\lambda + H_0$ is a closed convex set, there exists a unique $Y_0 \in X_\lambda + H_0$ such that

$$\sqrt{E\{Y_0^2\}} = \min\left\{\sqrt{E\{|X_\lambda + Y|^2\}} : Y \in H_0\right\} > 0.$$

Then clearly $Y_0 \perp H_0$. Since $X_\lambda - Y_0 \in H_0$, $X_\lambda - Y_0 \perp Y_0$ for all λ. This implies that Y_0 is independent of $\{X_\lambda - Y_0, \lambda \in \Lambda\}$. □

2. Proof of Theorem 1

As stated in Introduction let $X = \{X(x),\ x \in \mathrm{R}^d\}$ be a Brownian motion with d-dimensional time. For any fixed pair of real numbers t_1 and t_2 such that $0 < t_1 < t_2$ we put $\Lambda = \{x \in \mathrm{R}^d : t_1 \le |x| \le t_2\}$, $H = \text{c.l.s.}\{X(x),\ x \in \Lambda\}$ and $H_0 = \text{c.l.s.}\{X(x) - X(y), x, y \in \Lambda\}$. First we prepare the following lemma.

Lemma 2. *The condition* (B) *is satisfied for* $\{X(x),\ x \in \Lambda\}$, *namely, there exists* $x \in \Lambda$ *such that* $X(x) \notin H_0$.

Proof. (i) We consider the case where the dimension d is odd and $d \ge 3$. Denoting by $\hat{d}\theta$ the uniform distribution on $S^{d-1} = \{\theta \in \mathrm{R}^d : |\theta| = 1\}$, we put

$$\begin{aligned} R(t) &= \int_{S^{d-1}} X(t\theta)\hat{d}\theta, \qquad t \ge 0, \\ H_1 &= \text{c.l.s.}\{R(t), \quad t_1 \le t \le t_2\}, \\ H_1^\perp &= \text{the orthogonal complement of } H_1 \text{ in } H. \end{aligned}$$

Then we have

$$X(x) - R(|x|) \in H_1^\perp \qquad \text{for any } x \in \Lambda. \tag{2.1}$$

In fact, it is easy to see that, for each fixed $t \ge 0$, $E\{(X(x) - R(|x|))R(t)\}$ depends only on $|x|$ and hence it must vanish, which implies (2.1). We are going to prove that $X(t_1\theta) \notin H_0$ for $\theta \in S^{d-1}$. The relation (2.1) implies that $X(t_1\theta) = R(t_1) + X'$ with $X' \in H_1^\perp$ and that $H_0 \subset H_{10} \oplus H_1^\perp$ where $H_{10} = \text{c.l.s.}\{R(t) - R(s), t, s \in [t_1, t_2]\}$. Therefore, for the proof of $X(t_1\theta) \notin H_0$ it is enough to show that $R(t_1) \notin H_{10}$. We now make use of the canonical representation of the Gaussian process $\{R(t),\ t \ge 0\}$ due to McKean [5], which means that

$$R(t) = \int_0^t f(t, r)\, dB(r), \quad t \ge 0,$$

where $\{B(r),\ r \geq 0\}$ is a one-dimensional standard Brownian motion and

$$f(t,r) = k(d) \int_{r/t}^{1} (1-u^2)^{(d-3)/2}\, du, \qquad 0 \leq r \leq t, \tag{2.2}$$

$k(d)$ being a suitable constant depending only on d. For any s and t with $t_1 \leq s < t \leq t_2$ we have

$$\begin{aligned} R(t) - R(s) &= \int_0^{t_1} f_{ts}(r)\, dB(r) + \int_{t_1}^{t} g_{ts}(r) dB(r), \\ R(t_1) &= \int_0^{t_1} f(r)\, dB(r), \end{aligned}$$

where $f_{ts}(r) = f(t,r) - f(s,r)$, $f(r) = f(t_1, r)$ and $g_{ts}(r)$ is a suitable function. Therefore, if we put

$$\begin{aligned} \tilde{H}_0 &= \text{c.l.s.} \left\{ \int_0^{t_1} f_{ts}(r)\, dB(r), t, s \in [t_1, t_2] \right\}, \\ \tilde{H}_+ &= \text{c.l.s.} \{B(u) - B(r), r, u \in [t_1, t_2]\}, \end{aligned}$$

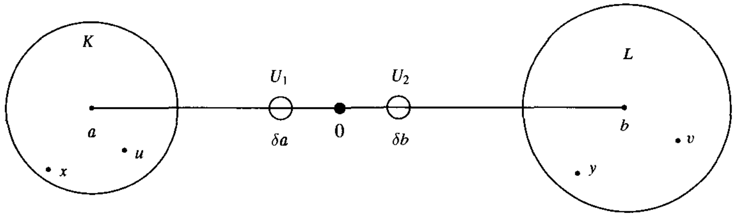
then $\tilde{H}_0 \perp \tilde{H}_+$, $H_{10} \subset \tilde{H}_0 \oplus \tilde{H}_+$ and $R(t_1) \perp \tilde{H}_+$. From these observations we see that for the proof of $R(t_1) \notin H_{10}$, it is enough to show

$$\int_0^{t_1} f(r) dB(r) \notin \tilde{H}_0. \tag{2.3}$$

Let L_0^2 be the subspace of $L^2[0, t_1]$ spanned by the functions $f_{ts}(\cdot), t, s \in [t_1, t_2]$. Then the Hilbert space $\tilde{H}_0$ is isomorphic to L_0^2 and (2.3) is equivalent to $f \notin L_0^2$. Now the assumption that d is an odd integer ≥ 3 implies that $f_{ts}(r), t, s, \in [t_1, t_2]$, are polynomials of degree $d-2$ vanishing at $r = 0$ (use (2.2)). Therefore all the functions in L_0^2 are also polynomials of degree at most $d - 2$ vanishing at $r = 0$. On the other hand it is easy to see that f is a polynomial of degree $d-2$ with $f(0) > 0$. Therefore $f \notin L_0^2$, which finally implies $X(t_1\theta) \notin H_0$. This completes the proof in the case where d is odd and $d \geq 3$.

(ii) The proof in the case where d is even can be obtained by the method of descent in which a Brownian motion with d-dimensional time is viewed as the restriction of a Brownian motion with $(d+1)$-dimensional time to $\mathrm{R}^d \times \{0\} \subset \mathrm{R}^{d+1}$ and also by using Remark 3. The proof in the case $d = 1$ is easy. The proof of Lemma 2 is finished. □

We are now able to prove Theorem 1. From the assumption on A_k, $1 \leq k \leq n$, there exist t_1 and t_2 with $0 < t_1 < t_2$ such that $\Lambda = \{x \in \mathrm{R}^d : t_1 \leq |x| \leq t_2\}$ includes all A_k. Then, by Lemma 2 the condition (B) is satisfied for $X_\Lambda = \{X(x), x \in \Lambda\}$



and by Remark 1 the condition (1.1) is satisfied for the function $f(w) = c_1 w(A_1) + c_2 w(A_1) + \cdots + c_n W(A_n)$, $w \in C(\Lambda)$, with $c = c_1 + \cdots + c_n$. Therefore by Lemma 1 the probability distribution of the random variable $f(X_\Lambda) = c_1 X(A_1) + c_2 X(A_2) + \cdots + c_n X(A_n)$ has a form $\gamma * \mu$. This completes the proof of Theorem 1.

3. Proof of Theorem 2

Under the assumption on A_j and B_k in Theorem 2 we can take disjoint closed balls K and L with the following properties:
(3.1) $K \supset \cup_{j=1}^{m} A_j$, $\quad L \supset \cup_{k=1}^{n} B_k$.
(3.2) K is separated from L by the hyperplane Π.
(3.3) The center a of K and the center b of L are on the straight line that passes through the origin 0 and is perpendicular to Π.
We consider open balls U_1 and U_2 with a common radius ε and with centers δa and δb, respectively, where $\delta > 0$ is chosen so that $\delta a \notin K$ and $\delta b \notin L$ (see the figure). We now make use of the Chentsov representation of $X(x)$ ([1]), which asserts that

$$X(x) = W(D_x), \tag{3.4}$$

where D_x is the open ball with center $x/2$ and radius $|x|/2$, and $\{W(d\xi)\}$ is a suitable white noise in R^d associated with the measure $c_d|\xi|^{-d+1}d\xi$ (c_d is a suitable constant), namely, a Gaussian random measure in R^d such that $E\{W(d\xi)\} = 0$ and $E\{W(d\xi)^2\} = c_d|\xi|^{-d+1}d\xi$. By taking $\varepsilon > 0$ small enough, we can assume

$$U_1 \subset \left\{\bigcap_{x\in K} D_x\right\} \cap \left\{\bigcup_{y\in L} D_y\right\}^c, \qquad U_2 \subset \left\{\bigcap_{y\in L} D_y\right\} \cap \left\{\bigcup_{x\in K} D_x\right\}^c. \tag{3.5}$$

If we write $X(x) = W(D_x) = W(U_1) + \tilde{X}_x$ and $X(y) = W(D_y) = W(U_2) + \tilde{X}_y$, then (3.5) implies that the 2-dimensional random vector $(W(U_1), W(U_2))$ is independent of the Gaussian family $\{(\tilde{X}_x, \tilde{X}_y) : x \in K, y \in L\}$. Therefore we have

$$f_1(X) = cW(U_1) + \tilde{f}_1, \qquad f_2(X) = c'W(U_2) + \tilde{f}_2,$$

with $c = \sum_{j=1}^{m} c_j$, $c' = \sum_{k=1}^{n} c'_k$ and $(W(U_1), W(U_2))$ is independent of $(\tilde{f}_1, \tilde{f}_2)$. Since $W(U_1)$ and $W(U_2)$ are independent and each of them is a nondegenerate Gaussian random variable with mean 0, the joint distribution of $f_1(X)$ and $f_2(X)$ has a form $(\gamma_1 \otimes \gamma_2) * \nu$.

4. Proof of Theorem 3 and Theorem 4

By using the fact that $\{X(x)-X(x_0), x \in \mathrm{R}^d\}$ is identical in law to $\{X(x-x_0), x \in \mathrm{R}^d\}$ for each $x_0 \in \mathrm{R}^d$ and also by using the assumption $\sum_{j=1}^{m} c_j = \sum_{k=1}^{n} c'_k$, we see that the probability distribution of $f_1(X) - f_2(X)$ is invariant under any simultaneous shift of A_j and B_k. Therefore, in proving Theorem 3 we may assume that A_j, and B_k satisfy the same assumption as in Theorem 2. Then the joint distribution of $f_1(X)$ and $f_2(X)$ has a form $(\gamma_1 \otimes \gamma_2) * \nu$ by Theorem 2 and this implies the conclusion of Theorem 3.

Before going to the proof of Theorem 4 we introduce some notation. Denote by $\mathcal{K}$ the set of all pairs (K_1, K_2) of disjoint closed balls K_1 and K_2 with rational centers and rational radii. We put $f(K_1, K_2; \sigma_1, \sigma_2) = X(K_1; \sigma_1) - X(K_2; \sigma_2)$ where each σ_i is either 0 or 1 and $X(K_i; \sigma_i)$ denotes either $\underline{X}(K_i)$ or $\overline{X}(K_i)$ according as $\sigma_i = 0$ or 1. We also denote by $\mathcal{E}(K_1, K_2; \sigma_1, \sigma_2)$ the event $\{f(K_1, K_2; \sigma_1, \sigma_2) = 0\}$ and then put $\mathcal{E}' = \cup \mathcal{E}(K_1, K_2; \sigma_1, \sigma_2)$ where the union is taken over all $(K_1, K_2) \in \mathcal{K}$ and all $(\sigma_1, \sigma_2) \in \{0, 1\}^2$. Finally let $\mathcal{E}$ be the event such that there exist distinct extreme-points x and y with $X(x) = X(y)$. It is then easy to see that $\mathcal{E} \subset \mathcal{E}'$. On the other hand Theorem 3 implies $P\{\mathcal{E}(K_1, K_2; \sigma_1, \sigma_2)\} = 0$ and hence $P\{\mathcal{E}'\} = 0$. This implies $P\{\mathcal{E}\} = 0$ as was to be proved.

5. Remarks on a diffusion process in a d-dimensional Brownian environment

This section is to supply an example for the applicability of Theorem 4. We change the notation for a Brownian motion with a d-dimensional time since we want to use $X(t)$ for a diffusion process. Let $\mathbf{W}$ be the space of continuous functions on R^d vanishing at 0. In this section an element W of $\mathbf{W}$ is called an environment. We consider the probability measure P on W such that $\{W(x),\ x \in \mathrm{R}^d, P\}$ is a Lévy's Brownian motion with a d-dimensional time. Let Ω be the space of continuous functions on $[0, \infty)$ taking values in R^d. The value of $\omega (\in \Omega)$ at time t is denoted by $X(t) = X(t, \omega) = \omega(t)$. For each fixed environment W we consider the probability measure P_W on Ω such that $\{X(t), t \geq 0, P_W\}$ is a diffusion process in R^d with generator

$$\frac{1}{2}(\Delta - \nabla W \cdot \nabla) = \frac{1}{2} e^W \sum_{k=1}^{d} \frac{\partial}{\partial x_k} \left(e^{-W} \frac{\partial}{\partial x_k} \right)$$

and starting from 0. Let $\mathcal{P}$ be the probability measure on $W \times \Omega$ defined by $\mathcal{P}(dWd\omega) = P(dW)P_W(d\omega)$. Then $\{X(t), t \geq 0, \mathcal{P}\}$ can be regarded as a process defined on the probability space $(W \times \Omega, \mathcal{P})$, which we call a diffusion process in a d-dimensional Brownian environment. When $d = 1$, this model is a diffusion analogue of well-known Sinai's random walk in a random environment(1982) and much is known about the long-term behavior of X(t) such as localization. When $d \geq 2$, a similar diffusion model appeared in [3]. Now our interest is the long-term behavior of $\{X(t), t \geq 0, \mathcal{P}\}$ in the case $d \geq 2$. Tanaka [6](see also [7]) proved that, for any dimension d, $\{X(t), t \geq 0, P_W\}$ is *recurrent* for almost all Brownian sample environments W. Mathieu[4] proved that *localization* takes place for $\{X(t), t \geq 0, \mathcal{P}\}$, in the sense that

$$\lim_{N\to\infty} \overline{\lim_{\lambda\to\infty}} \mathcal{P}\left(\lambda^{-2}\max\{|X(t)| : 0 \leq t \leq e^{\lambda}\} > N\right) = 0.$$

However, in the case $d \geq 2$, it seems that the existence of the limiting distribution of $\{\lambda^{-2}X(e^{\lambda}), \mathcal{P}\}$ as $\lambda \to \infty$ is still an open problem. We give a remark on this problem. We notice the scaling relation

$$\{X(t),\ t \geq 0,\ P_{\lambda W_\lambda}\} \stackrel{d}{=} \{\lambda^{-2}X(\lambda^4 t),\ t \geq 0,\ P_W\},$$

where $\lambda > 0$ and $W \in \mathbf{W}$ are fixed, W_λ denotes an element of W defined by $W_\lambda(x) = \lambda^{-1}W(\lambda^2 x)$, $x \in \mathrm{R}^d$, and $\stackrel{d}{=}$ means the equality in distribution. This scaling relation combined with $W_\lambda \stackrel{d}{=} W$ imply the following: If we can prove that $\{X(e^{r\lambda}),\ P_{\lambda W}\}$ has the limiting distribution as $\lambda \to \infty$ under the condition $r = r(\lambda) \to 1$, then so does $\{\lambda^{-2}X(e^{\lambda}), \mathcal{P}\}$. From now on we are interested in $\{X(t), P_{\lambda W}\}$. For $W \in \mathbf{W}$ we define the sub-level domain D as the connected component of the open set $\{x \in \mathrm{R}^d : W(x) < 1\}$ containing 0. Then it is easy to see that D is bounded, P-a.s. *By making use of* Theorem 4 we see that for W not belonging to some P-negligible subset of $\mathbf{W}$, there exists a point $\tilde{b}$ of local (strict) minimum of W with depth > 1 inside D. Such a point $\tilde{b}$ is characterized by (i) $W(\tilde{b}) < W(x)$ for $x \in U - \{\tilde{b}\}$ and (ii) $U \subset D$, where U denotes the connected component of the open set $\{x \in \mathrm{R}^d : W(x) - W(\tilde{b}) < 1\}$ containing $\tilde{b}$. It is obvious that the totality of such points $\tilde{b}$ is a finite set, which is denoted by $\{b_k(W),\ 1 \leq k \leq l(W)\}$. Now suppose $l(W) = 1$ and put $b = b_1(W)$. Then from the argument of [4] we see that

$$X(e^{r\lambda}) \to b \text{ (in probability with respect to } P_{\lambda W}) \tag{5.1}$$

as $\lambda \to \infty$ provided $r = r(\lambda)$(non-random) tends to 1. If $l(W) \geq 2$, we do not know whether the limiting distribution of $X(e^{r\lambda})$ exists. Hoping for the best, we think it might be possible to define b, in one way or another, as *a single point* among $b_k(W)$, $1 \leq k \leq l(W)$, and to prove (5.1) even in the case $l(W) \geq 2$, for almost all w.

References

[1] N.N. Chentsov: *Lévy Brownian motion for several parameters and generalized white noise*, Theory Probab. Appl. **2** (1957), 265–266 (English translation).

[2] P. Lévy: Processus Stochastiques et Mouvement Brownien, Gauthier-Villars, Paris, 1948.

[3] E. Marinari, G. Parisi, D. Ruelle and P. Windy: *On the interpretation of* $1/f$ *noise*, Comm. Math. Phys. **89** (1983), 1–12.

[4] P. Mathieu: *Zero white noise limit through Dirichet forms, with application to diffusions in a random medium*, Probab. Theory Related Fields, **99** (1994), 549–580.

[5] H.P. McKean: *Brownian motion with a several-dimensional time*, Theory Probab. Appl. **8** (1963), 335–354.

[6] H. Tanaka: *Recurrence of a diffusion process in a multidimensional Brownian environment*, Proc. Japan Acad. Math. Sci. **69** (1993), 377–381.

[7] H. Tanaka: *Diffusion processes in random environments*, Proc. International Congress of Mathematicians, Zürich, 1994, Birkhäuser, Basel, 1995, 1047–1054.

Department of Mathematical and Physical Sciences
Japan Women's University
2-8-1 Mejirodai Bunkyo-ku, Tokyo 112-8681
Japan

Present address: 1-4-17-104 Miyamaedaira
Miyamae-ku, Kawasaki 216-0006
Japan

Bibliography of Hiroshi Tanaka

All of $*$ are contained in this volume.

[1] On limiting distributions for one-dimensional diffusion processes, Bull. Math. Statist., **7** (1957), 84-91.

[2] Some properties of one-dimensional diffusion processes (with G. Maruyama), Mem. Fac. Sci., Kyushu Univ., Ser. A, **11** (1957), 117–141.

[3] Certain limit theorems concerning one-dimensional diffusion processes, Mem. Fac. Sci., Kyushu Univ., Ser. A, **12** (1958), 1–11.

[4] Ergodic property of N-dimensional recurrent Markov processes (with G. Maruyama), Mem. Fac. Sci., Kyushu Univ., Ser. A, **13** (1959), 157–172.

[5]* Additive functionals of the Brownian path (with H.P. McKean), Mem. Coll. Sci., Kyoto Univ., Ser. A, **33** (1961), 479–506.

[6] Local times on the boundary for multi-dimensional diffusion (with K. Sato), Proc. Japan Acad., **38** (1962), 699–702.

[7]* Note on continuous additive functionals of the 1-dimensional Brownian path, Z. Wahrscheinlichkeitstheorie verw. Gebiete, **1** (1963), 251–257.

[8]* Existence of diffusions with continuous coefficients, Mem. Fac. Sci., Kyushu Univ., Ser. A, **18** (1964), 89–103.

[9] Note on additive functionals of the class (U) of Meyer of Markov processes, Mem. Fac. Sci., Kyushu Univ., Ser. A, **18** (1964), 154–165.

[10] Local solutions of stochastic differential equations associated with certain quasilinear parabolic equations, J. Fac. Sci., Univ. Tokyo, Sec. I, **14** (1967), 313–326.

[11] On the pre-closedness of the potential operator (with K. Yosida and T. Watanabe), J. Math. Soc. Japan, **20** (1968), 419–421.

[12] Propagation of chaos for certain Markov processes of jump type with nonlinear generators I, II, Proc. Japan Acad., **45** (1969), 449-452; 598–600.

[13] Purely discontinuous Markov processes with nonlinear generators and their propagation of chaos, Teor. Verojatnost. Primenen., **15** (1970), 599–621.

[14]* Propagation of chaos for certain purely discontinuous Markov processes with interactions, J. Fac. Sci., Univ. Tokyo, Sec. I, **17** (1970), 259–272.

[15]* An inequality for a functional of probability distributions and its application to Kac's one-dimensional model of a Maxwellian gas, Z. Wahrscheinlichkeitstheorie verw. Gebiete, **27** (1973), 47–52.

[16]* On Markov process corresponding to Boltzmann's equation of Maxwellian gas, Proc. Second Japan-USSR Symp. on Probab. Th., 478–489, Lecture Notes in Math., **330**, Springer-Verlag, 1973.

[17] An inequality for certain functional of multidimensional probability distributions (with H. Murata), Hiroshima Math. J., **4** (1974), 75–81.

[18] Perturbation of drift-type for Lévy processes (with M. Tsuchiya and S. Watanabe), J. Math. Kyoto Univ., **14** (1974), 73–92.

[19]* On the uniqueness of Markov process associated with the Boltzmann equation of Maxwellian molecules, Proc. Intern. Symp. on Stochastic Differential Equations, 1976, 409–425, Kinokuniya, 1978.

[20] On the growing up problem for semilinear heat equations (with K. Kobayashi and T. Sirao), J. Math. Soc. Japan, **29** (1977), 407–424.

[21] Homogenization of certain one-dimensional discontinuous Markov processes (with M. Horie and T. Inuzuka), Hiroshima Math. J., **7** (1977), 629–641.

[22]* Probabilistic treatment of the Boltzmann equation of Maxwellian molecules, Z. Wahrscheinlichkeitstheorie verw. Gebiete, **46** (1978), 67–105.

[23] Stochastic differential equation associated with the Boltzmann equation of Maxwellian molecules in several dimensions, Stochastic Analysis, Proc. Intern. Conf. held at Northwestern Univ., 1978, 301–314, Academic Press, 1978.

[24]* Stochastic differential equations with reflecting boundary condition in convex regions, Hiroshima Math. J., **9** (1979), 163–177.

[25] Central limit theorem for a simple diffusion model of interacting particles (with M. Hitsuda), Hiroshima Math. J., **11** (1981), 415–423.

[26] Fluctuation theory for Kac's one-dimensional model of Maxwellian molecules, Sankhyā, Ser. A, Part I, **44** (1982), 23–46.

[27]* Some probabilistic problems in the spatially homogeneous Boltzmann equation, Theory and Application of Random Fields, 258–267, Lecture Notes in Control and Information Sciences, **49**, Springer-Verlag, 1983.

[28] Limit distribution of the minimum distance between independent and identically distributed d-dimensional random variables (with T. Onoyama and M. Sibuya), Statistical Extremes and Applications, 549–562, Reidel, 1984.

[29]* Limit theorem for certain diffusion processes with interaction, Stochastic Analysis : Proceedings of the Taniguchi Intern. Symp. on Stochastic Analysis, Katata and Kyoto, 1982, 469–488, Kinokuniya, 1984.

[30] Homogenization of diffusion processes with boundary conditions, Advances in Probab., 411–437, Marcel Dekker, 1984.

[31] A diffusion process in a singular mean-drift-field (with M. Nagasawa), Z. Wahrscheinlichkeitstheorie verw. Gebiete, **68** (1985), 247–269.

[32]* Central limit theorem for a system of Markovian particles with mean field interaction (with T. Shiga), Z. Wahrscheinlichkeitstheorie verw. Gebiete, **69** (1985), 439–459.

[33]* Propagation of chaos for diffusing particles of two types with singular mean field interaction (with M. Nagasawa), Probab. Th. Rel. Fields, **71** (1986), 69–83.

[34]* Stochastic differential equations for mutually reflecting Brownian balls (with Y. Saisho), Osaka J. Math., **23** (1986), 725–740.

[35]* Limit distribution for 1-dimensional diffusion in a reflected Brownian medium, Seminaire de Probabilites XXI, 246–261, Lecture Notes in Math., **1247**, Springer-Verlag, 1987.

[36]* Limit distributions for one-dimensional diffusion processes in self-similar random environments, Hydrodynamic Behavior and Interacting Particle Systems, 189–210, IMA Volumes, **9**, Springer-Verlag, 1987.

[37] On the symmetry of a reflecting Brownian motion defined by Skorohod's equation for multi-dimensional domain (with Y. Saisho), Tokyo J. Math., **10** (1987), 419–435.

[38]* Stochastic differential equation corresponding to the spatially homogeneous Boltzmann equation of Maxwellian and non-cutoff type, J. Fac. Sci., Univ. Tokyo, Sec. IA, **34** (1987), 351–369.

[39] Diffusion with interaction and collisions between coloured particles and the propagation of chaos (with M. Nagasawa), Probab. Th. Rel. Fields, **74** (1987), 161–198.

[40] One-dimensional diffusions and random walks in random environments (with K. Kawazu and Y. Tamura), Probability Theory and Mathematical Statistics, Fifth Japan-USSR Symposium Proceedings, 1986, 170–184, Lecture Notes in Math., **1299**, Springer-Verlag, 1988.

[41]* Limit theorem for one-dimensional diffusion process in Brownian environment, Stochastic Analysis, 156–172, Lecture Notes in Math., **1322**, Springer-Verlag, 1988.

[42] On the propagation of chaos for diffusion processes with drift coefficients not of average form (with M. Nagasawa), Tokyo J. Math., **10** (1989), 403–418.

[43] Time reversal of random walks in one-dimension, Tokyo J. Math., **12** (1989), 159–174.

[44] A theorem of Pitman type for simple random walks on $\boldsymbol{Z}^d$ (with H. Miyazaki), Tokyo J. Math., **12** (1989), 235–240.

[45] Limit theorems for one-dimensional diffusions and random walks in random environments (with K. Kawazu and Y. Tamura), Probab. Th. Rel. Fields, **80** (1989), 501–541.

[46] Time reversal of random walks in $\boldsymbol{R}^d$ Tokyo J. Math., **13** (1990), 375–389.

[47] Localization of diffusion processes in one-dimensional random environment (with K. Kawazu and Y. Tamura), J. Math. Soc. Japan, **44** (1992), 515–550.

[48] Limit theorems of the minimum interpoint distance between any pair of i.i.d. random points in $\boldsymbol{R}^d$ (with S. Kanagawa and Y. Mochizuki), Ann. Inst. Statist. Math., **44** (1992), 121–131.

[49]* On the maximum of a diffusion process in a drifted Brownian environment (with K. Kawazu), Seminaire de Probabilites XXVII, 78–85, Lecture Notes in Math., **1557**, Springer-Verlag, 1993.

[50] Green operators of absorbing Lévy processes on the half line, Stochastic Processes, A Festschrift in honour of G. Kallianpur, 313–319, Springer-Verlag, 1993.

[51]* Recurrence of a diffusion process in a multidimensional Brownian environment, Proc. Japan Acad., **69**, Ser. A (1993), 377–381.

[52]* Localization of a diffusion process in a one-dimensional Brownian environment, Comm. Pure Appl. Math., **47** (1994), 755–766.

[53] Localization of one-dimensional random walks in random environment (with K. Kawazu and Y. Tamura), Proc. Intern. Conf. on "Stochastic Processes, Physics and Geometry II", 1991, 444–452, World Scientific, 1995.

[54]* Diffusion processes in random environments, Proc. Intern. Congress of Mathematicians, Zürich, Switzerland 1994, 1047–1054, Birkhäuser, 1995.

[55]* Environment-wise central limit theorem for a diffusion in a Brownian environment with large drift, Itô's Stochastic Calculus and Probability Theory, 373–384, Springer-Verlag, 1996.

[56]* A diffusion process in a Brownian environment with drift (with K. Kawazu), J. Math. Soc. Japan, **49** (1997), 189–211.

[57]* Limit theorems for a Brownian motion with drift in a white noise environment, Chaos, Solitons & Fractals, **8** (1997), 1807–1816.

[58]* Invariance principle for a Brownian motion with large drift in a white noise environment (with K. Kawazu), Hiroshima Math. J., **28** (1998), 129–137.

[59] Stochastic differential equations of pure-jumps in relativistic quantum theory (with M. Nagasawa), Chaos, Solitons & Fractals, **10** (1999), 1265–1280.

[60] Time dependent subordination and Markov processes with jumps (with M. Nagasawa), Seminaire de Probabilités XXXIV, 257-288, Lecture Notes in Math., **1729**, Springer-Verlag, 2000.

[61] The principle of variation for relativistic quantum particles (with M. Nagasawa), Seminaire de Probabilités XXXV, 1-27, Lecture Notes in Math., **1755**, Springer-Verlag, 2001.

[62] A diffusion process with a one-sided Brownian potential (with K. Kawazu, and Y. Suzuki), Tokyo J. Math., **24** (2001), 211-229.

[63]* Some theorems concerning extrema of Brownian motion with d-dimensional time, Osaka J. Math., **38** (2001), 369-377.

Permissions

The editors and World Scientific Publishing Co. Pte. Ltd. would like to thank the original publishers of Tanaka's papers for granting permissions to reprint specific papers in this volume. The following list contains the credit lines for those articles.

[5] Reprint from Mem. Coll. Sci., Kyoto Univ., Ser. A, **12**, © 1961 by Kyoto University.

[7] Reprint from Z. Wahrscheilichkeitstheorie verw. Gebiete, **1**, © 1963 by Springer-Verlag.

[8] Reprint from Mem. Fac. Sci., Kyushu Univ., Ser. A, **18**, © 1964 by Kyushu University.

[14] Reprint from J. Fac. Sci., Univ. Tokyo, Sec. I, **17**, © 1970 by the University of Tokyo.

[15] Reprint from Z. Wahrscheilichkeitstheorie verw. Gebiete, **27**, © 1973 by Springer-Verlag.

[16] Reprint from Lecture Notes in Mathematics, **330**, © 1973 by Springer-Verlag.

[19] Reprint from Proc. Intern. Symp. on Stochastic Differential Equations, © 1978 by Kinokuniya Co. Ltd.

[22] Reprint from Z. Wahrscheilichkeitstheorie verw. Gebiete, **46**, © 1978 by Springer-Verlag.

[24] Reprint from Hiroshima Math. J., **9**, © 1979 by Hiroshima University.

[27] Reprint from Lecture Note in Control and Information Sciences, **49**, © 1983 by Springer-Verlag.

[29] Reprint from Stochastic Analysis : Proceedings of the Taniguchi Intern. Symp. on Stochastic Analysis, Katata and Kyoto, 1982, © 1984 by Kinokuniya Co. Ltd.

[32] Reprint from Z. Wahrscheilichkeitstheorie verw. Gebiete, **69**, © 1985 by Springer-Verlag.

[33] Reprint from Probab. Th. Rel. Fields, **71**, © 1986 by Springer-Verlag.

[34] Reprint from Osaka J. Math., **23**, © 1986 by Osaka University.

[35] Reprint from Lecture Notes in Mathematics, **1247**, © 1987 by Springer-Verlag.

[36] Reprint from Hydrodynamic Behavior and Interacting Particle Systems, (ed. G. Papanicolaou), IMA Volumes in Mathematics and Its Applications, **9**, © 1987 by Springer-Verlag.

[38] Reprint from J. Fac. Sci., Univ. Tokyo, Sec. IA, **34**, © 1987 by the University of Tokyo.

STOCHASTIC PROCESSES

Selected Papers of Hiroshi Tanaka

Hiroshi Tanaka is noted for his discovery of the "Tanaka formula", which is a generalization of the Itô formula in stochastic analysis. This important book is a selection of his brilliant works on stochastic processes and related topics. It contains Tanaka's papers on (i) Brownian motion and stochastic differential equations (additive functionals of Brownian paths and stochastic differential equations with reflecting boundaries), (ii) the probabilistic treatment of nonlinear equations (Boltzmann equation, propagation of chaos and McKean-Vlasov limit), and (iii) stochastic processes in random environments (especially limit theorems on the stochastic processes in one-dimensional random environments and their refinements). The book also includes essays by Henry McKean, Marc Yor, Shinzo Watanabe and Hiroshi Tanaka on Tanaka's works.

ISBN 981-02-4591-2

9 789810 245917

www.worldscientific.com

4673 hc