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

θ = the colatitude of x'

φ = 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:

$$\frac{\pi-\theta}{2} = \int_0^{\rho_0} \frac{d\rho}{\sqrt{1-\rho^2-\dfrac{4}{|x-x_1|^2}\,U\!\left(\dfrac{r}{\rho}\right)}}, \tag{0.1}$$

0044-3719/78/0046/0067/$07.80

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,

$$X(t)=X(0)+\sum_{s\leq t} a(X(s-), Y(s-), p(s)); \tag{0.5b}$$

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 $\sigma\{$——$\}$ denotes the smallest σ-field on Ω that makes $\{$——$\}$ 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

$$\begin{aligned} &E\{\exp(i\zeta \sum_{\tau < s \leqq t} A(s, p(s), \omega)) | \mathscr{F}_\tau\} \\ &\quad = \exp\{\int_{(\tau, t] \times S} (e^{i\zeta A(s, \sigma, \omega)} - 1)\, ds\, \lambda(d\sigma)\}, \qquad \zeta \in \mathbf{R}. \end{aligned} \tag{2.2}$$

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

$$\begin{aligned}&|a(x,x,\theta,\varphi)-a(y,y_1,\theta,\varphi+\varphi_0)|\\&=|a(y,y_1,\theta,\varphi)|\leqq\frac{|y-y_1|}{2}\theta\\&\leqq\tfrac{1}{2}\{|x-y|+|x_1-y_1|\}\,\theta.\end{aligned}$$

(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

$$\begin{aligned}\tau z&=\frac{|y-y_1|}{|x-x_1|}\left(z-\frac{x+x_1}{2}\right)+\frac{x+x_1}{2},\\\tilde{\tau}z&=z+\frac{y+y_1}{2}-\frac{x+x_1}{2}.\end{aligned}$$

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

$$|a(x,x_1,\theta,\varphi)-(\rho A(x,x_1,\theta,\varphi)-x^*)| \\ \leqq|a(x,x_1,\theta,\varphi)|\times(\text{the rotation angle of }\rho) \\ \leqq\frac{\pi}{2}|x-x^*|\,\theta, \tag{3.2}$$

$$|(\rho A(x,x_1,\theta,\varphi)-x^*)-(\tau\rho A(x,x_1,\theta,\varphi)-\tau x^*)| \\ \leqq\tfrac{1}{2}\{|x-x_1|-|y-y_1|\}\,\theta\leqq\tfrac{1}{2}\{|x-y|+|x_1-y_1\}\,\theta, \tag{3.3}$$

$$\begin{aligned}\tau\rho A(x,x_1,\theta,\varphi)-\tau x^* &=\tilde{\tau}\tau\rho A(x,x_1,\theta,\varphi)-\tilde{\tau}\tau x^* \\ &=A(y,y_1,\theta,\varphi+\varphi_0(x,x_1,y,y_1))-y \\ &=a(y,y_1,\theta,\varphi+\varphi_0(x,x_1,y,y_1)).\end{aligned} \tag{3.4}$$

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:

$$\text{For each fixed } w\in W,\ P(w,\cdot) \text{ is a probability measure on } W. \tag{3.5}$$

$$\text{For each fixed } A\in\mathscr{B}_W,\ P(\cdot, A) \text{ is a Borel function on } W. \tag{3.6}$$

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 $Y_0, \ldots, Y_{n-1}$ *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* $Y_0, \ldots, Y_{n-1}$).

(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

$$\begin{aligned} &E\{e^{i\zeta\cdot X_\Delta(t)}|\mathscr{F}_{t_k}\}\\ &=\exp\{i\zeta\cdot X_\Delta(t_k)+(t-t_k)\int_S(e^{i\zeta\cdot a(X_\Delta(t_k), y, \theta, \varphi)}-1)\,u_k(dy)\,Q(d\theta)\,d\varphi\}\\ &\qquad\qquad t_k\leqq t\leqq t_{k+1},\ \zeta\in\mathbf{R}^3, \end{aligned}$$

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*

$$\begin{aligned} &X_\Delta(0)=X \\ &X_\Delta(t)=X_\Delta(t_k)+\sum_{t_k<s\leqq t} a(X_\Delta(t_k), Y_k, p(s)), \qquad t_k<t\leqq t_{k+1}. \end{aligned} \tag{4.9}$$

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*

$$\begin{aligned} &E|\mathring{X}(t)-\mathring{X}^{\#}(t)| \\ &\quad\leqq \{1+c_1(t-t_k)\}\, E|\mathring{X}(t_k)-\mathring{X}^{\#}(t_k)|+c_1(t-t_k)\,\rho_1(u_k, u_k^{\#}), \\ &\qquad\qquad t_k<t\leqq t_{k+1} \quad (0\leqq k<n), \end{aligned} \tag{4.10}$$

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* $\square$: $0=s_0<s_1<\cdots<s_m=T$ *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* $\{X_\square(t), 0\leqq t\leqq T\}$, *respectively, and satisfying*

$$E|X(t)-\tilde{X}(t)|\leqq c_2(|\Delta|+|\square|), \qquad 0\leqq t\leqq T, \tag{4.12}$$

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 $\square$ 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 $Y_0,\ldots,Y_{n-1}$, 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 $\{X_\square(t), 0\leqq t\leqq T\}$.

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

$$\begin{aligned}
&E|X(t)-\tilde{X}(t)|\\
&\quad\leqq E|X(s_k)-\tilde{X}(s_k)|+E\Big\{\int_{(s_k,t]\times S} |a(X(t_{k'}), Y_{k'}, \sigma)-\tilde{a}(\tilde{X}(s_k), \tilde{Y}_k, \sigma)|\,ds\,\lambda(d\sigma)\Big\}\\
&\quad\leqq E|X(s_k)-\tilde{X}(s_k)|+(t-s_k)\,2\pi c\int_0^\pi \theta Q(d\theta)\{E|X(t_{k'})-\tilde{X}(s_k)|+E_\alpha|Y_{k'}-\tilde{Y}_k|\}\\
&\quad=E|X(s_k)-\tilde{X}(s_k)|+c_0(t-s_k)\,E|X(t_{k'})-\tilde{X}(s_k)|,
\end{aligned} \tag{4.15}$$

where $c_0 = 4\pi c \int_0^\pi \theta Q(d\theta)$. On the other and

$$\begin{aligned} E|X(t)| &\leqq E|X(t_k)| + (t-t_k) E \int_S |a(X(t_k), Y_k, \sigma)| \lambda(d\sigma) \\ &\leqq E|X(t_k)| + (t-t_k) E \int_S (|X(t_k) - Y_k|/2) \theta \lambda(d\sigma) \\ &\leqq E|X(t_k)| + c'(t-t_k) E|X(t_k)|, \qquad t_k < t \leqq t_{k+1}, \end{aligned}$$

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

$$\begin{aligned} E|X(s_k) - X(t_{k'})| &\leqq c'(s_k - t_{k'}) E|X(t_{k'})| \\ &\leqq c''(s_k - t_{k'}), \qquad c'' = E|X| c' e^{c'T}, \end{aligned}$$

and hence

$$\begin{aligned} &E|X(t) - \tilde{X}(t)| + c''|\Delta| \\ &\quad \leqq \{E|X(s_k) - \tilde{X}(s_k)| + c''|\Delta|\} e^{c_0(t-s_k)}, \qquad s_k < t \leqq s_{k+1}, \end{aligned}$$

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 $t_k<t\leqq t_{k+1}$

$$\begin{aligned} &E|X^{\Delta}(t)-X^{\#}(t)| \\ &\quad\leqq E|X^{\Delta}(t_k)-X^{\#}(t_k)|+c_1(t-t_k)E|X(t_k)-X^{\#}(t_k)| \\ &\quad\leqq \{1+c_1(t-t_k)\}E|X^{\Delta}(t_k)-X^{\#}(t_k)|+c_1(t-t_k)E|X(t_k)-X^{\Delta}(t_k)|, \end{aligned} \tag{4.20}$$

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

$$E|X^{\Delta}(t)-X^{\#}(t)|\leqq\varepsilon(\Delta)(e^{c_1 T}-1), \tag{4.21}$$

$$E|X(t)-X^{\#}(t)|\leqq\varepsilon(\Delta)e^{c_1 T}. \tag{4.22}$$

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

$$\begin{aligned} \rho_1(u_k, u_k^{\#}) &\leqq E|\mathring{X}(t_k)-\mathring{X}^{\#}(t_k)|+E|X^{\#}(t_k)-X(t_k)| \\ &\leqq E|\mathring{X}(t_k)-\mathring{X}^{\#}(t_k)|+\varepsilon(\Delta)e^{c_1 T}. \end{aligned}$$

Therefore, (4.10) yields

$$\begin{aligned} E|\mathring{X}(t)-\mathring{X}^{\#}(t)| &\leqq \{1+c_0(t-t_k)\}E|\mathring{X}(t_k)-\mathring{X}^{\#}(t_k)| \\ &\qquad +c_1(t-t_k)\varepsilon(\Delta)e^{c_1 T}, \qquad t_k<t\leqq t_{k+1}, \end{aligned}$$

which implies that

$$E|\mathring{X}(t)-\mathring{X}^{\#}(t)|\leqq\varepsilon(\Delta)e^{c_1 T}(e^{c_0 T}-1), \qquad 0\leqq t\leqq T. \tag{4.23}$$

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

$$\begin{aligned}\tilde{E}|\tilde{X}(t)-\tilde{X}_\Delta(t)| &\leqq E|X(t)-X^{\#}(t)|+E|\mathring{X}^{\#}(t)-\mathring{X}(t)| \\ &\leqq \varepsilon(\Delta)\, e^{3c_1 T}, \quad 0 \leqq t \leqq T. \end{aligned} \tag{4.24}$$

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

$$\tilde{E}|\tilde{X}(t)-\tilde{X}_\Delta(t)| \leqq c_3 |\Delta| + \varepsilon_Y(\Delta) e^{c_1 T}, \quad 0 \leqq t \leqq T, \tag{4.25}$$

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*

$$\tilde{E}|\tilde{X}(t)-\tilde{X}^{\#}(t)| \leqq e^{c_0 t} \rho_1(f, f^{\#}), \quad t \geqq 0. \tag{4.26}$$

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

there exists an α-random variable which is independent of Y and uniformly distributed on (0.1). (4.27)

We then define $\{X_1(t)\}$ by

$$X_1(t) = X + \sum_{s \leq t} a_1(X_0(s-), Y_0(s-), p(s)),$$

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 $\{X_k(t)\}$, $1 \leq k \leq n$, are defined together with auxiliary α-processes $\{Y_k(t)\}$ and $\{\mathscr{F}_t\}$-predictable processes $\{\varphi_k^*(t, \alpha, \omega)\}$, $0 \leq k < n$, 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

$$\sum_{j=1}^{\nu} \frac{1}{j!} \sum_{\substack{\nu_1 + \dots + \nu_j = \nu \\ \nu_1, \dots, \nu_j \geqq 1}} \frac{\nu!}{\nu_1! \dots \nu_j!} \prod_{l=1}^{j} ((t - t_k) \int_S |a(X_\Delta(t_k), Y_k, \sigma)|^{\nu_l} \lambda(d\sigma)).$$

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 $t_k < t \leqq t_{k+1}$

$$\begin{aligned} &E\{|X_\Delta(t)|^\nu\} \\ &\quad \leqq E\{|X_\Delta(t_k)|^\nu\} + \bar{c}_\nu \int_{(t_k, t] \times (0, 1)} E\{|X_\Delta(s)|^\nu + |X_\Delta(t_k)|^\nu + |Y_k|^\nu\} ds\, d\alpha \\ &\quad = \{1 + 2\bar{c}_\nu (t - t_k)\} E\{|X_\Delta(t_k)|^\nu\} + \bar{c}_\nu \int_{t_k}^{t} E\{|X_\Delta(s)|^\nu\} ds, \end{aligned}$$

$$\bar{c}_\nu = 3^{\nu-1} \nu\pi \int_0^\pi \theta Q(d\theta).$$

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 $\mathscr{B}_t=\sigma\{X_s\colon s\leqq t\}$, $\mathscr{B}=\vee\,\mathscr{B}_t(=\mathscr{B}_W)$. 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

$$\tilde{E}|\tilde{X}(t)-\tilde{X}^{\#}(t)|\leqq e^{c_1 t}\{|x-y|+e^{c_0 t}\rho_1(f,g)\}, \qquad t\geqq 0.$$

This assertion immediately implies the following lemma, in which we put

$$e_f(t,x,A)=P_f^x\{X_t\in A\}, \qquad A\in\mathscr{B}(R^3). \tag{5.2}$$

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*

$$P_f^x\{X_{t_1}\in A\,|\,\mathscr{B}_{t_0}\}=e_{u(t_0)}(t_1-t_0, X_{t_0}, A), \qquad P_f^x\text{-a.s.}, \tag{5.3}$$

$$P_f\{X_{t_1}\in A\,|\,\mathscr{B}_{t_0}\}=e_{u(t_0)}(t_1-t_0, X_{t_0}, A), \qquad P_f\text{-a.s.}, \tag{5.4}$$

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

$$p^{\#}(t)=p(t_0+t), \qquad t>0,$$
$$\mathscr{F}_0^{\#}=\{\phi,\Omega\}, \qquad \mathscr{F}_t^{\#}=\sigma\{p^{\#}(s),\ 0<s\leqq t\},$$

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

$$dX^{\#}(t)=a(X^{\#}(t-), Y^{\#}(t-), \varphi+\varphi^{\#})\,dN^{\#}, \qquad X^{\#}(0)=y, \tag{5.5}$$

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 $t_1 > t_0$

$$\begin{aligned}
&P\{X^0(t_1)\in A\,|\,\mathscr{F}_{t_0}\}\\
&\quad = P\{X^{\#}_{X(t_0)}(t_1 - t_0)\in A\,|\,\mathscr{F}_{t_0}\}\\
&\quad = P\{X_y^{\#}(t_1 - t_0)\in A\}, \qquad y = X(t_0), \text{ a.s.}\\
&\quad = P_{u(t_0)}^{X(t_0)}\{X_{t-t_0}\in A\} = e_{u(t_0)}(t - t_0, X(t_0), A), \qquad \text{a.s.}
\end{aligned} \tag{5.6}$$

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 $\mathfrak{e}$ 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

$$\mathfrak{E}(F) = \int_{R^3\times R^3} |x-y|^2\, F(dx\,dy),$$

$$\mathfrak{e}(f, g) = \inf_{F\in\mathbf{F}(f,g)} \mathfrak{E}(F), \qquad \rho(f, g) = \sqrt{\mathfrak{e}(f, g)},$$

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 $\mathfrak{e}(f, g)$ 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 $\mathfrak{e}(f, g) \leq \lim_{n\to\infty} \mathfrak{e}(f_n, g_n)$.

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

$$\begin{aligned}&\text{if } F\in\mathbf{F}(f,g) \text{ and } \mathfrak{E}(F)=\mathfrak{e}(f,g), \text{ then}\\ &F(A\times B)=\int_B \delta_{\psi(x)}(A)\,g(dx) \text{ for any } A,B\in\mathscr{B}(R^3) \text{ with a}\\ &\text{suitable Borel mapping } \psi \text{ from } \mathbf{R}^3 \text{ into itself,}\end{aligned}\tag{6.2}$$

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

$$\mathfrak{e}(f,g)\leq E\{\mathfrak{e}(f^\omega,g^\omega)\}.\tag{6.3}$$

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 $\varepsilon\downarrow 0$, $\lim_{N\to\infty}\sup_{0<\varepsilon<1}\int_{|x|>N}|x|^2 f_\varepsilon(dx)=0$, 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, $F_\varepsilon=\int_\Omega F_\varepsilon^\omega P(d\omega)$ clearly belongs to $\mathbf{F}(f_\varepsilon,g_\varepsilon)$ and hence we have

$$\mathfrak{e}(f_\varepsilon,g_\varepsilon)\leqq\mathfrak{E}(F_\varepsilon)=\int_\Omega\mathfrak{E}(F_\varepsilon^\omega)\,P(d\omega)=E\{\mathfrak{e}(f_\varepsilon^\omega,g_\varepsilon^\omega)\}.$$

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,

$$\mathfrak{e}(U_{x,r,l},U_{y,s,m})\leqq E\{|X-Y|^2\}=\frac{1}{2\pi}\int_0^{2\pi}|X(\omega)-Y(\omega)|^2\,d\omega.$$

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 $Y_0^\varepsilon, \ldots, Y_{n-1}^\varepsilon$ *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

$$\bar{\mathrm{e}}_\theta(f, g) \geqq 0, \tag{7.3}$$

$$\mathrm{e}(T_t f, T_t g) \leqq \mathrm{e}(f, g) - 2\pi \int_0^t ds \int_0^\pi \bar{\mathrm{e}}_\theta(T_s f, T_s g)\, Q(d\theta). \tag{7.4}$$

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.

$$\mathrm{e}(\Pi_{x, x_1, \theta}, \Pi_{y, y_1, \theta}) \leqq \Phi_\theta(x, x_1, y, y_1),$$

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:

$$\begin{aligned} &x \to \{x+x_1+(x-x_1)\cos\theta\}/2, && y \to \{y+y_1+(y-y_1)\cos\theta\}/2,\\ &r \to |x-x_1|(\sin\theta)/2, && s \to |y-y_1|(\sin\theta)/2,\\ &l \to (x-x_1)/|x-x_1|, && m \to (y-y_1)/|y-y_1|. \end{aligned}$$

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) $\mathfrak{e}[(f_1 \circ f_2)_\theta, (g_1 \circ g_2)_\theta] \leqq \frac{1+\cos\theta}{2}\mathfrak{e}(f_1, g_1) + \frac{1-\cos\theta}{2}\mathfrak{e}(f_2, g_2), \qquad \theta \in (0, \pi).$

(ii) $\mathfrak{e}(f_1 \circ f_2, g_1 \circ g_2) \leqq \gamma\,\mathfrak{e}(f_1, g_1) + (1-\gamma)\,\mathfrak{e}(f_2, g_2),$ *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) $E\{|X_1-Y_1|^2\} = \mathfrak{e}(f_1, g_1), \quad E\{|X_2-Y_2|^2\} = \mathfrak{e}(f_2, g_2).$

(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

$$\begin{aligned} &\mathfrak{e}[(f_1 \circ f_2)_\theta, (g_1 \circ g_2)_\theta]\\ &\quad \leqq E\{\mathfrak{e}(\Pi_{X_1, X_2, \theta}, \Pi_{Y_1, Y_2, \theta})\} \leqq E\{\Phi_\theta(X_1, X_2, Y_1, Y_2)\}\\ &\quad = E\left\{\left|\frac{1+\cos\theta}{2}(X_1-Y_1) + \frac{1-\cos\theta}{2}(X_2-Y_2)\right|^2\right\}\\ &\quad\quad + \frac{\sin^2\theta}{4}E\{|X_1-X_2|^2 + |Y_1-Y_2|^2 - |X_1-X_2||Y_1-Y_2|\\ &\quad\quad\quad - |(X_1-X_2, Y_1-Y_2)|\}. \end{aligned}$$

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

$$\begin{aligned}
&\mathrm{e}[(f_1\circ f_2)_\theta,(g_1\circ g_2)_\theta]\\
&\quad\leqq E\left\{\left|\frac{1+\cos\theta}{2}(X_1-Y_1)+\frac{1-\cos\theta}{2}(X_2-Y_2)\right|^2\right\}\\
&\qquad+\frac{\sin^2\theta}{4}E\{|X_1-X_2|^2+|Y_1-Y_2|^2-2(X_1-X_2,Y_1-Y_2)\}\\
&\quad=\frac{(1+\cos\theta)^2}{4}E\{|X_1-Y_1|^2\}+\frac{(1-\cos\theta)^2}{4}E\{|X_2-Y_2|^2\}\\
&\qquad+\frac{\sin^2\theta}{4}E\{|X_1-Y_1|^2\}+\frac{\sin^2\theta}{4}E\{|X_2-Y_2|^2\}\\
&\quad=\frac{1+\cos\theta}{2}\mathrm{e}(f_1,g_1)+\frac{1-\cos\theta}{2}\mathrm{e}(f_2,g_2).
\end{aligned}$$

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

$$\begin{aligned}&\varlimsup_{s\downarrow 0}\{\mathrm{e}(T_{t+s}f, T_{t+s}g)-\mathrm{e}(T_tf, T_tg)\}/s\\&\quad\leqq -q\{\mathrm{e}(T_tf, T_tg)-\mathrm{e}[(T_tf)^{(2)},(T_tg)^{(2)}]\}\\&\quad\leqq -q(2\pi/q)\int_0^\pi\{\mathrm{e}(T_tf, T_tg)-\mathrm{e}[[T_tf]_\theta,[T_tg]_\theta]\}\,Q(d\theta)\\&\quad=-2\pi\int_0^\pi\bar{\mathrm{e}}_\theta(T_tf, T_tg)\,Q(d\theta).\end{aligned}$$

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

$$\xi_{\mathbf{n}}(x)=|x|^{2r}\,\xi_k^l(x) \quad \text{for } \mathbf{n}=(r,k,l),$$
$$\xi_{(0,0,0)}(x)=1,$$

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 $\mu_{\mathbf{n}}(t)=\mu_{\mathbf{n}}(T_t f)$ for $\mathbf{n}$ such that $|\mathbf{n}|\leqq\nu$. Then, for any $\mathbf{n}$ with $|\mathbf{n}|\leqq\nu$ we have*

$$\frac{d}{dt}\mu_{\mathbf{n}}(t)=\sum{}'\,\beta^{\mathbf{n}}_{\mathbf{n}_1,\mathbf{n}_2}\,\mu_{\mathbf{n}_1}(t)\,\mu_{\mathbf{n}_2}(t)-\beta_{\mathbf{n}}\mu_{\mathbf{n}}(t), \tag{8.3}$$

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

$$\begin{aligned}\frac{d}{dt}\mu_{\mathbf{n}}(t) &= \langle (T_t f)\otimes(T_t f), K_{\mathbf{n}}(x,y)\rangle \\ &= \langle (T_t f)\otimes(T_t f), J_{\mathbf{n}}(x,y)\rangle \\ &\quad + \langle (T_t f)\otimes(T_t f), K_{\mathbf{n}}(x,0) + K_{\mathbf{n}}(0,y)\rangle \\ &= \sum' \beta^{\mathbf{n}}_{\mathbf{n}_1,\mathbf{n}_2}\,\mu_{\mathbf{n}_1}(t)\,\mu_{\mathbf{n}_2}(t) - \beta_{\mathbf{n}}\,\mu_{\mathbf{n}}(t),\end{aligned}$$

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*

$$\int |x-m|^2 f(dx) = 3v > 0, \qquad m = \int x f(dx). \tag{8.5}$$

Let $\mathfrak{g}$ 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 $0 \leqq |\mathbf{n}| < k$ 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

$$\mathfrak{e}[(f \circ f)_\theta, (\mathfrak{g} \circ \mathfrak{g})_\theta] \leqq \mathfrak{e}(f, \mathfrak{g}). \tag{9.1}$$

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,

$$\begin{aligned}&\frac{|\psi(y_1)-\psi(y_0)|}{|y_1-y_0|}(y_1-y_0)-\frac{|\psi(y_2)-\psi(y_0)|}{|y_2-y_0|}(y_2-y_0)\\&\quad=\psi(y_1)-\psi(y_2)=\frac{|\psi(y_1)-\psi(y_2)|}{|y_1-y_2|}(y_1-y_2), \quad \text{a.e.,}\end{aligned}$$

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 $\mathfrak{e}(T_t f, T_t \mathfrak{g})=\mathfrak{e}(T_t f, \mathfrak{g})=\mathfrak{e}(T_t f)$, the decreasing property of $\mathfrak{e}(T_t f)$ in t follows from Theorem 7.1. To prove that $\mathfrak{e}(T_t f)$ 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 $\int |x|^4 (T_t f)(dx)$ 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 $|\bar{\mathfrak{e}}_\theta(\tilde{f})-\bar{\mathfrak{e}}_\theta(\tilde{\mathfrak{g}})|\leqq 2\mathfrak{e}(\tilde{f},\tilde{\mathfrak{g}})$ for $\tilde{f},\tilde{\mathfrak{g}}\in\tilde{\mathscr{P}}_\varepsilon$ 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

$$\mathrm{e}(T_t f) \leqq \mathrm{e}(f) - 2\pi \int_0^t ds \int_0^\pi \bar{\mathrm{e}}_\theta(T_s f)\, Q(d\theta).$$

Because $T_t f \in \tilde{\mathscr{P}}_\varepsilon$ with $\varepsilon = \mathrm{e}(T_t f)$ if $\mathrm{e}(T_t f) > 0$, the above inequality combined with (9.4) implies

$$\mathrm{e}(T_t f) \leqq \mathrm{e}(f) - \int_0^t \Phi[\mathrm{e}(T_s f)]\, ds \tag{9.5}$$

for t such that $\mathrm{e}(T_t f) > 0$. But, this inequality clearly implies that $\mathrm{e}(T_t f) \to 0$ 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

$$\begin{aligned}\mathrm{e}(T_t f) &\leqq \{\rho(T_t f, T_t f_\varepsilon) + \rho(T_t f_\varepsilon, \mathfrak{g})\}^2 \\ &\leqq \{\varepsilon + \sqrt{\mathrm{e}(T_t f_\varepsilon)}\}^2\end{aligned}$$

and hence $\overline{\lim_{t\to\infty}}\, \mathrm{e}(T_t f) \leqq \varepsilon^2$. 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 $\mathrm{e}(T_t f) \to 0$ 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

$$\begin{aligned}&\int_{(0,\pi)\times(0,2\pi)\times\mathbf{R}^6} \xi(x)(u' u_1' - u u_1)\, Q(d\theta)\, d\varphi\, dx\, dx_1 \\ &\quad = \int_{\mathbf{R}^6} (K\xi)(x, x_1)\, u(x)\, u(x_1)\, dx\, dx_1, \qquad \xi \in C_0^\infty(\mathbf{R}^3),\end{aligned}$$

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 $\mathfrak{g}$ be the Gaussian distribution (8.6). Then $\mathfrak{g}$ 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

$$\langle \mathfrak{g}\otimes\mathfrak{g},K\xi\rangle=0,\qquad \xi\in C_0^\infty(\mathbf{R}^3).$$

This implies $T_t \mathfrak{g} = \mathfrak{g}$ 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 $T_t \mathfrak{g} = \lim_{\varepsilon \downarrow 0} T_t^{(\varepsilon)} \mathfrak{g} = \mathfrak{g}$, proving the invariance of $\mathfrak{g}$ under T_t. Moreover, $(\mathfrak{g} \circ \mathfrak{g})_\theta = \mathfrak{g}$ also follows from $g' g_1' = g g_1$ and the above lemma.

References

1. Arkeryd, L.: On the Bolztmann equation, Part I: Existence, Part II: The full initial value problem. Arch Rational Mech. Anal. **45**, 1–16, 17–34 (1972)
2. Carleman, T.: Problèmes Mathématiques dans la Théorie Cinétique des Gaz. Uppsala 1957.
3. Erdélyi, A. (Editor): Higher Transcendental Functions, Vol. II, 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 I. 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