APPLIED MATHEMATICS AND STATISTICS LABORATORIES

STANFORD UNIVERSITY
CALIFORNIA

CLASSICAL DIFFUSION PROCESSES AND TOTAL POSITIVITY

By
SAMUEL KARLIN and JAMES McGREGOR

TECHNICAL REPORT No. 1
January 20, 1960

PREPARED UNDER THE AUSPICES OF
NATIONAL SCIENCE FOUNDATION GRANTS G-9669 AND G-6332
CLASSICAL DIFFUSION PROCESSES AND TOTAL POSITIVITY

by

SAMUEL KARLIN and JAMES MCGREGOR

TECHNICAL REPORT NO. 1

January 20, 1960

PREPARED UNDER THE AUSPICES

OF

NATIONAL SCIENCE FOUNDATION GRANTS G-9669 and G-6332

APPLIED MATHEMATICS AND STATISTICS LABORATORIES

STANFORD UNIVERSITY

STANFORD, CALIFORNIA
CLASSICAL DIFFUSION PROCESSES AND TOTAL POSITIVITY

by

Samuel Karlin and James McGregor

Introduction

We take a diffusion equation

\[ Au - \frac{1}{\rho(x)} \frac{\partial}{\partial x} \left[ \kappa(x) \frac{\partial u}{\partial x} \right] = \frac{\partial u}{\partial t} \quad [u = u(x,t)] \]

on an interval \( a_1 < x < a_2 \), where \( \rho(x) \) and \( 1/\kappa(x) \) are measurable functions that are almost everywhere positive on \( (a_1,a_2) \) and integrable over all compact subintervals. In the simplest case \( \rho \) and \( 1/\kappa \) are integrable over the full interval \( (a_1,a_2) \), in which situation we take a pair of "classical" boundary conditions

\[ c_{1,1} u(a_1,t) + c_{1,2} \left( \kappa \frac{\partial u}{\partial x} \right)(a_1,t) = 0 \quad (i = 1,2). \]

For an essentially arbitrary function \( f \), equation (1) with conditions (2) has a solution with initial values \( u(x,0) = f(x) \) which is unique subject to natural and well-known continuity requirements. The solution may be written in the form

* This work has been done under individual National Science Foundation grants to the authors.
\[ u(x,t) = \int_{a_1}^{a_2} p(t;x,y)f(y)\rho(y)dy, \]

(3)

\[ p(t;x,y) = \sum_{n=0}^{\infty} e^{-\lambda_n^t} \varphi_n(x) \varphi_n(y), \]

where \( \{\lambda_n\} \) are the eigenvalues of the Sturm-Liouville problem

\[ A\varphi + \lambda \varphi = 0, \]

\[ c_{1,1}\varphi(a_1) + c_{1,2}(\kappa\varphi')(a_1) = 0 \quad (i = 1, 2) \]

and \( \{\varphi_n\} \) are the corresponding eigenfunctions, taken to be orthonormal in the sense that

\[ \int_{a_1}^{a_2} \varphi_m(x)\varphi_n(x)\rho(x)dx = \delta_{m,n}. \]

If the two boundary conditions are of dissipative type, i.e., if

\[ 0 \leq c_{11}/c_{12} \leq +\infty \quad \text{and} \quad -\infty \leq c_{21}/c_{22} \leq 0, \]

then the function

\[ P(t;x,\mathcal{E}) = \int_{\mathcal{E}} p(t;x,y)\rho(y)dy \]

is the transition probability function of a stationary Markoff process \( X(t) \) in the open interval \((a_1, a_2)\); more precisely,

\[ P(t;x,\mathcal{E}) = \Pr\{X(t+s) \in \mathcal{E} | X(s) = x\}. \]
We say that the process takes place in the open interval because the Chapman-Kolmogoroff equation

\[ P(t+s;x,E) = \int_{(a_1,a_2)} P(t;x,dy)P(s;y,E) \]

is valid when we integrate over only the open interval. Of course, to get a separable version of the process it may be necessary to adjoin one or both of the boundary points to the state space, and in fact, when the boundaries are regular as described above, this is always necessary.

If \( c_{1,2} = c_{2,2} = 0 \), then both of the boundaries are absorbing barriers and it is known (see Ray [14] and also the discussion in Feller [5]) that the corresponding process \( X(t) \) has continuous path functions. For the other boundary conditions (reflecting barriers and elastic barriers) it is known that the path functions are continuous in the interior of the interval, but their behavior at the boundaries has not yet been analyzed.

In this paper we have two objectives. We first develop a number of important examples of diffusion processes that possess a representation of the form (3) whose terms are classical orthogonal functions. Included in this list are the Hermite, Laguerre, and Jacobi systems of polynomials and Bessel functions. We find that these classical diffusion processes are virtually all derivable by suitable geometrical constructions from the Brownian motion process. More specifically, we start with the Brownian motion process in \( n \) dimensions in which the particle is constrained to move on a prescribed surface of revolution. We then consider the induced process whose motion is represented by the projection
of the Brownian motion on the axis of revolution. For suitable surfaces
of revolution the resulting one-dimensional process has a transition
density that can usually be expressed in terms of classical orthogonal
functions. We also note that these same examples of classical diffusion
processes frequently arise as idealizations of stochastic models in
biology and physics.

Our second objective is to show that the transition probability
function for these examples has, for every $t > 0$, the property of total
positivity. The analytic property of total positivity is intimately
related to the stochastic property that the path functions of the process
are continuous everywhere, including at the boundaries [9]. The precise
description of this property is as follows.

Given two Borel sets $E,F$ in $(a_1,a_2)$, the notation $E < F$ will
signify that $x < y$ for every $x \in E$ and every $y \in F$. We take for
any positive integer $n$ a collection of points $x_1 < x_2 < \ldots < x_n$ and
a collection of Borel sets $E_1 < E_2 < \ldots < E_n$, and form the determinant

$$(4) \quad P(t;x_1,E_1;x_2,E_2;\ldots;x_n,E_n) = \det P(t;x_i,E_j).$$

With fixed $t$ the function $P(t;\cdot;\cdot)$ is called positive of order $m$ if
all such determinants for $n = 1,\ldots,m$ are non-negative, and totally
positive if it is positive of every finite order. Alternatively, we may
take two sets of $n$ points, $x_1 < x_2 < \ldots < x_n$ and $y_1 < y_2 < \ldots < y_n$,
and with the density function form the determinant
\[ p \left( t; x_1, x_2, \ldots, x_n \right) = \det p(t; x_1, y_j) ; \]

and if these determinants are non-negative we say that \( p(t; \cdot, \cdot) \) is positive of order \( m \) or totally positive as the case may be. The two kinds of determinants are related by the formula

\[
\Phi \left( t; x_1, \ldots, x_n \right) = \int_{E_1} \cdots \int_{E_n} p \left( t; y_1, \ldots, y_n \right) \rho(y_1) \cdots \rho(y_n) \, dy_1 \cdots dy_n .
\]

The continuity properties of the path functions of the process give rise to the positivity of (4) by virtue of the following result of the authors [9], referred to later as the coincidence probability theorem.

Suppose we have a stationary strong Markoff process \( X(t) \) on a (possibly closed) interval \( I \) of the real line, with path functions continuous everywhere on \( I \), including at the boundaries. In \( I \) let \( x_1 < \ldots < x_n \) be \( n \) points and \( E_1 < \ldots < E_n \) be \( n \) Borel sets. We take \( n \) labeled particles and allow them to execute the \( X(t) \) process simultaneously and independently. If the first particle starts at \( x_1 \), ..., and the \( n^{th} \) particle at \( x_n \), then the determinant (4) is the probability that at time \( t \) the first particle will be in the set \( E_1 \), ..., and the \( n^{th} \) particle in the set \( E_n \), without any two of the particles ever having been simultaneously at the same place in the intervening time. (For the truth of this assertion we must know that the direct product of \( n \) copies of the \( X(t) \) process has the strong Markoff property. For this technical detail, and for the exact meaning of the picturesque language, the reader is referred to [9].)
Thus we see that, roughly speaking, continuity everywhere of the path functions implies total positivity of the transition probability function. Now although the behavior of the path functions at the boundaries for the elastic barrier processes is not known, there are strong reasons for expecting that they are continuous there. In fact, for the analogous discrete-state diffusion process, the so-called birth and death processes, the path functions are known to have the appropriate continuity properties [3]. This is the main reason for expecting that the function (3) is totally positive.

In a subsequent paper we shall study in detail the total positivity properties of various fundamental solutions of (1) both in the case of regular boundaries and when one or both boundaries are not regular. In this paper we confine ourselves to establishing total positivity by special arguments for a number of important examples.

In general, there are two lines of reasoning. On the one hand, it may be possible to express the function \( p(t;x,y) \) simply in terms of well-known functions and deduce total positivity from this representation. This kind of argument is non-probabilistic and comparatively elementary. On the other hand, it may be possible to start from some well-known process such as Brownian motion in space and by a suitable geometrical construction derive a one-dimensional process governed by an equation of the same kind as (1). If this geometrical construction makes it evident that the path functions of the derived process are everywhere continuous, then we assume that the coincidence probability theorem applies and consider the total positivity to be established. This latter type of
argument is not intended to be rigorous and is presented only briefly with no discussion of the strong Markoff property. It serves primarily as background for the analytic arguments to follow.

In Section 1 we examine the radial process of an n-dimensional Brownian motion, and present an explicit expression of the transition probability function in terms of Bessel functions. This formula is essentially known. Since the spectrum is continuous, the representation is an integral formula rather than an infinite sum. We use this formula to prove total positivity in a direct manner. An alternative method of proof, based on the coincidence probability theorem, exploits the familiar continuity properties of the Brownian motion.

We also consider the radial process of the Brownian motion under the condition that the particle never touches the surface of the unit sphere. The spectrum is now discrete, and we secure a representation of the form (3). The eigenfunctions again involve Bessel functions.

In Section 2 we start with a brief discussion of the well-known Ornstein-Uhlenbeck process. This diffusion process has a transition density which is expressed in terms of Hermite polynomials. Here, the state of the process can be regarded as the velocity of a free particle attracted to the origin by a force directly proportional to its distance from the origin. If we now consider the induced random variable equal to the distance from the origin at time \( t \), we obtain a diffusion process whose transition density has a representation involving the associated Laguerre polynomial of order \(-1/2\). This follows from the identity
\( H_{2m}(x) = (-1)^m 2^m m! L_m^{1/2}(x^2) \),

where \( H \) and \( L \) designate the standard Hermite and Laguerre polynomials, respectively.

This example may be generalized as follows. Suppose a particle executing a Brownian motion in \( n \) dimensions is subjected to a restoring force directed to the origin and directly proportional to the distance from the origin. We consider the induced radial process of this motion. In Section 2 it is shown that the transition density function can be exhibited as a sum of Laguerre polynomials and exponential factors. In each of these examples the property of total positivity is verified by using the classical bilinear generating function satisfied by the Laguerre polynomials.

These same diffusion processes have been studied by Feller [4] as models of population growth. Feller's aim was to determine the most general solution of the forward equation of the process. Our interest, by contrast, is in displaying explicit representations of the transition density function in terms of classical orthogonal functions, and in verifying the total positivity of this density function.

In Sections 3 and 4 we study diffusion processes whose representation formulas involve Jacobi polynomials. We begin with the Brownian motion process taking place on the sphere in \( n \) dimensions. Then we construct the one-dimensional process whose state variable is the orthogonal projection of the motion on a fixed diameter through the origin. The transition density of this diffusion process is expressible in terms of Jacobi
polynomials. Unfortunately, the approach by means of geometrical construction on the Brownian motion leads to the Jacobi diffusion process only for special cases of its parameters. In Section 4 we obtain the general Jacobi diffusion process as a limiting case of a birth and death process involving the Hahn polynomial system.

These approximations in terms of Hahn polynomials, while of independent interest, are also useful in establishing the property of total positivity.

The Jacobi diffusion process also occurs as a stochastic model of gene frequencies (see also [11], [19]).

1. The radial process of Brownian motion

   If \( Y(t) = (Y_1(t), \ldots, Y_N(t)) \) is the \( N \)-dimensional Brownian motion process with diffusion equation \( \partial u / \partial t = \Delta u \) and if

   \[
   X(t) = \left[ \sum_{i=1}^{N} Y_i^2(t) \right]^{1/2}
   \]

   is its "radial part," then \( X(t) \) is a one-dimensional process on \( 0 < x < \infty \).

   A path function of the \( X \) process is observed by starting a sample of the \( Y \) process and recording the distance from the origin. The \( Y \) process has continuous path functions, and this property is evidently inherited by the \( X \) process. The diffusion equation of the \( X \) process is known ([12], [15]) to be

   \[
   \frac{\partial u}{\partial t} = \frac{1}{x^{N-1}} \frac{\partial}{\partial x} \left( x^{N-1} \frac{\partial u}{\partial x} \right) = \frac{\partial^2 u}{\partial x^2} + \frac{2\gamma}{x} \frac{\partial u}{\partial x} \quad (0 < x < \infty),
   \]
where \( \gamma = (N-1)/2 \) and the appropriate boundary condition at \( x = 0 \) is

\[
\left( x^{N-1} \frac{\partial u}{\partial x} \right)_{x=0} = 0.
\]

The boundary at \( \infty \) is not regular, and it transpires that no boundary condition is required there. The transition probability function is

\[
P(t;x,E) = \int_E p(t;x,y) d\mu(y),
\]

where

\[
\mu(y) = 2^{-(\gamma+1/2)} \left[ \Gamma(\gamma + \frac{3}{2}) \right]^{-1} y^{2\gamma+1}
\]

and

\[
p(t;x,y) = \int_0^\infty e^{-\alpha^2 t} T(\alpha x) T(\alpha y) d\mu(\alpha),
\]

\[
T(x) = \Gamma(\gamma + \frac{1}{2}) \left( \frac{x}{\alpha^2} \right)^{1/2} J_{\gamma-1/2}(x).
\]

The notation here is that of Bochner [1].

On the basis of the geometrical construction argument we conclude by appealing to the coincidence probability theorem that (5) is totally positive. Let us now turn to a special analytical argument which will prove that (5) is totally positive not merely when \( \gamma = (N-1)/2 \) \((N = 1, 2, \ldots, )\), but for arbitrary real \( \gamma > -1/2 \). An alternative method for computing the transition density of the \( X \) process is to express the transition density of the \( Y \) process in spherical polar coordinates, multiply by the element of Lebesgue area on the unit sphere \( \Omega \), and integrate over \( \Omega \) with respect
to the final state variable. This gives

\[ p(t;x,y) = (2t)^{-(\gamma + 1/2)} e^{x^2/4t} e^{-y^2/4t} T(i\frac{tx}{2t}) \]

with \( T(x) \) as before. Now \( T(ix/2t) \) is a power series in \( x^2 \) with positive coefficients.

\[ T(i\frac{tx}{2t}) = \sum_{n=0}^{\infty} a_n x^{2n}, \]

all \( a_n > 0 \). Hence, if \( 0 < x_1 < \ldots < x_m \) and \( 0 < y_1 < \ldots < y_m \), then

\[ \det T(i\frac{tx_\alpha y_\beta}{2t}) = \prod_{0 \leq n \leq m} a_n \cdot \det(x_\alpha^n) \det(y_\beta^n), \]

which is positive, since all the Vandermonde determinants are positive; and from this follows the total positivity of \( p(t;x,y) \). The two functions (5) and (6) are indeed the same for arbitrary real \( \gamma > -1/2 \); see [17, p. 395].

The geometrical argument is capable of several interesting variations. For example, if \( Y(t) \) denotes the \( N \)-dimensional Wiener process as before, and \( Z(t) \) denotes the \( Y(t) \) process started inside the unit sphere and terminated whenever it first reaches the surface of the unit sphere, then we may observe \( X_1(t) \), the distance of \( Z(t) \) from the origin. We see that \( X_1(t) \) is a Markoff process on \( 0 < x \leq 1 \) with continuous path functions for which the point \( 1 \) is an absorbing state. Its (backward) diffusion equation is

\[ \frac{\partial u}{\partial t} = \frac{1}{x^{N-1}} \frac{\partial}{\partial x} \left( x^{N-1} \frac{\partial u}{\partial x} \right), \]
with the boundary conditions \( u = 0 \) at \( x = 1 \) and \( x^{N-1} \frac{\partial u}{\partial x} = 0 \) at \( x = 0 \). The transition function is
\[
P(t;x_0,E) = \int_E p(t;x_0,y) y^{N-1} dy,
\]
where
\[
p(t;x_0,y) = \sum_{n=0}^{\infty} e^{-\lambda_n^2 t} \varphi_n(x) \varphi_n(y) \psi_n,
\]
with
\[
\varphi_n(x) = x^{-(N-2)/2} J_{(N-2)/2}(x \sqrt{\lambda_n}), \quad \psi_n = \left[ \int_0^1 \varphi_n^2(x) x^{N-1} dx \right]^{-1},
\]
and \( \{ \lambda_n^2 \} \) the sequence of positive zeros of \( J_{(N-2)/2}(x) \). Thus by appealing to the coincidence probability theorem we conclude that (8) is totally positive.

In a similar way it is possible to construct a Markov process on \( 1 \leq x < \infty \) with continuous path functions, (7) as its diffusion equation, and an absorbing barrier at \( x = 1 \).

There does not appear to be a simple formula for (8) analogous to (6).

2. Brownian motion of a harmonically bound particle

The diffusion equation
\[
\frac{\partial u}{\partial t} = x^2 \frac{\partial}{\partial x} \left( e^{-x^2} \frac{\partial u}{\partial x} \right) = \frac{\partial^2 u}{\partial x^2} - 2x \frac{\partial u}{\partial x}
\]
on \( -\infty < x < \infty \) is the backward equation for a particle executing a Brownian motion and attracted to the origin by a restoring force directly proportional to its distance from the origin. There is an alternative interpretation (the Ornstein-Uhlenbeck process) also connected with the
Brownian motion. It is known that the process has continuous path functions. We conclude that the transition density

\begin{equation}
    p(t;x,y) = \sum_{n=0}^{\infty} e^{-2nt} H_n(x) H_n(y) w_n,
\end{equation}

where

\begin{equation}
    H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}
\end{equation}

are the Hermite polynomials and \( w_n = \left(\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{n!}{n!}\right)^{-1} \), is a totally positive function.

By using the bilinear generating function for Hermite polynomials [2, formula 22, p. 194] we may write (9) in the form

\begin{equation}
    p(t;x,y) = \pi^{-1/2} (1-e^{-4t})^{-1/2} \exp\left(\frac{-x^2 e^{-4t}}{1-e^{-4t}}\right) \exp\left(\frac{-y^2 e^{-4t}}{1-e^{-4t}}\right) \exp\left(\frac{2xy e^{-2t}}{1-e^{-4t}}\right).
\end{equation}

The density function of the process with respect to Lebesgue measure reduces to the familiar formula

\begin{equation}
    p(t;x,y) e^{-y^2} = \frac{1}{\sqrt{\pi} \sqrt{1-e^{-4t}}} \exp\left(-\frac{(xe^{-2t} - y)^2}{1-e^{-4t}}\right).
\end{equation}

We recognize this as a Brownian motion with a mean position \( xe^{-2t} \) at time \( t \) (\( x \) is the initial position) and variance \( (1-e^{-4t})/2 \). Now for fixed \( t > 0 \) the function

\begin{equation}
    f(x) = \exp\left(\frac{2xe^{-2t}}{1-e^{-4t}}\right)
\end{equation}
is a power series in $x$ with positive coefficients. It follows that $f(xy)$ is totally positive in $x,y$, and hence that $p(t;x,y)$ is totally positive.

Next we consider a Brownian motion in $N$ dimensions with a restoring force proportional to its distance from the origin. The diffusion equation

$$\frac{\partial u}{\partial t} = \Delta u - 2r \frac{\partial u}{\partial r}.$$  

The fundamental solution of this equation is unique, and the corresponding process undoubtedly has continuous path functions, although we can cite no reference to this effect. (This is not relevant to our proof of total positivity, since our argument is analytic and does not depend on the continuity of path functions.) If for this $N$-dimensional process we observe $X(t)$, the distance from the origin, then $X(t)$ is a Markoff process (the restoring force is spherically symmetric) on $0 < x < \infty$; this process should also have continuous path functions, and its transition function should be totally positive. Now the backward diffusion equation of the $X$ process is

$$\frac{\partial u}{\partial t} = \frac{1}{x^{N-1}} \frac{\partial}{\partial x} \left( x^{N-1} \frac{\partial u}{\partial x} \right) - 2x \frac{\partial u}{\partial x}$$

$$= \frac{\partial^2 u}{\partial x^2} + \left( \frac{N-1}{x} - 2x \right) \frac{\partial u}{\partial x} \quad (0 < x < \infty),$$

with the boundary condition $x^{N-1} \frac{\partial u}{\partial x} = 0$ at $x = 0$. No condition at
\( \infty \) is required, and the transition function is

\[
P(t; x, E) = \int_E \rho(t; x, y) y^{N-1} e^{-y^2} dy,
\]

where

\[
p(t; x, y) = \sum_{n=0}^{\infty} e^{-nt} L_n^{(\alpha)}(x^2) L_n^{(\alpha)}(y^2) w_n,
\]

\( L_n^{(\alpha)}(x) \) being the Laguerre polynomials with parameter \( \alpha = (N-2)/2 \) and normalizing constants

\[
w_n^{-1} = \frac{1}{2} \frac{\Gamma(\alpha+1)}{\Gamma(\frac{n+\alpha}{2})}.
\]

(The notation \( L_n^{(\alpha)} \) is that of Szegö [16].) Now the bilinear generating function for the Laguerre polynomials [2, formula 20, p. 189] gives us an alternative formula for (10), namely

\[
p(t; x, y) = \frac{2}{(1 - e^{-4t})} \exp \left( \frac{-x^2 e^{-4t}}{1 - e^{-4t}} \right) \exp \left( \frac{-y^2 e^{-4t}}{1 - e^{-4t}} \right) f(xy),
\]

where \( f \) is the Bessel function

\[
f(x) = (xe^{-2t})^{-\alpha} L_\alpha \left( 2 \frac{xe^{-2t}}{1 - e^{-4t}} \right),
\]

which is a power series with positive coefficients. The identity of (10) and (11) is valid for any real \( \alpha > -1 \). Thus we see, by analogy with (6), that (10) is actually totally positive for any real \( \alpha > -1 \).

A class of diffusion processes which also involve Laguerre polynomials comes from studies of population growth [3]. We consider a forward equation of the form
where $a$, $b$, and $c$ are constants and $a > 0$. Here, the random variable $x(t)$ denotes the population size at time $t$; $bx + c$ the instantaneous mean rate of growth when the population size is $x$, and $ax$ the instantaneous variance.

Feller studied (12) and determined the most general probability solution under certain conditions. He exhibited explicit solutions for certain choices of the parameters $a$, $b$, and $c$. We are interested in the symmetric representations of the form (3) involving Laguerre polynomials. Our results are supplementary to those of Feller and should be compared with his.

By separating the variables, we compute the transition density and obtain

\begin{equation}
    p(t;x,\xi) = \sum_{n=0}^{\infty} \frac{e^{-(n+1)bt}}{n!} \omega_n \text{L}_n^{(\alpha)}(\frac{b}{a}x) \text{L}_n^{(\alpha)}(\frac{b}{a}\xi),
\end{equation}

where $\alpha = 1 - \frac{c}{a}$ and $\text{L}_n^{(\alpha)}(\cdot)$ denote the Laguerre polynomials of order $\alpha$ as in (10). Here,

\begin{equation}
    \omega_n^{-1} = \int_{0}^{\infty} [\text{L}_n^{(\alpha)}(\xi)]^2 d\psi(\xi), \quad d\psi(\xi) = \frac{\alpha^{-\xi} e^{-\xi}}{\Gamma(\alpha+1)} d\xi.
\end{equation}

The transition probability function is

\begin{equation}
    P(t;x,E) = \int_{E} p(t;x,\xi)d\xi(\xi),
\end{equation}

where

\begin{equation}
    d\xi(\xi) = \frac{e^{-\xi/a}(b/a)^{\alpha+1} \alpha}{\Gamma(2-(c/a))} d\xi.
\end{equation}
Equation (13) is valid only for $b > 0$ and $2 > c/a$.

Using the bilinear generating relation [2, p. 189], we obtain

\[
(14) \quad p(t; x, \xi) = \frac{b}{a(e^{bt} - 1)} \exp \left\{ \frac{-b}{a} \left( x + \xi e^{bt} \right) \right\} \left( \frac{x}{\xi} e^{-bt} \right)^{-\alpha/2} I_{\alpha} \left[ \frac{2b \sqrt{x \xi} e^{-bt}}{a(1 - e^{-bt})} \right],
\]

which agrees with Feller's formula. Here $I_{\alpha}$ denotes the Bessel function in standard notation. From (14) we deduce in the usual manner that $p(t; x, \xi)$ is totally positive.

If instead of (12) we consider its formal adjoint, the backward equation

\[
u_t = a\nu_{xx} + (bx + c)\nu_x \quad (0 < x < \infty),
\]

we get

\[
(15) \quad p(t; x, \xi) = \sum_{n=0}^{\infty} e^{nt} \frac{\omega L_n(\alpha)}{\omega L_n(\alpha)} \left( \frac{a}{b} x \right)^{-\alpha/2} I_{\alpha} \left[ \frac{2 \sqrt{b/a} \xi e^{-bt}}{1 - e^{-bt}} \right],
\]

which is valid for $\alpha = c/a - 1$, $b < 0$, and $c > 0$.

An alternative formula is

\[
p(t; x, y) = \frac{-b}{a(1 - e^{-bt})} \exp \left\{ \frac{b(x + \xi e^{bt})}{a(1 - e^{-bt})} \right\} \left( \frac{x}{\xi} e^{-bt} \right)^{-\alpha/2} I_{\alpha} \left[ \frac{2 \sqrt{(b^2/a^2) \xi e^{-bt}}}{1 - e^{-bt}} \right].
\]

Equations (13) and (15) are the only cases in which we get solutions of (12) and (14), respectively, exhibited as infinite sums of Laguerre polynomials.

3. Jacobi polynomials and Brownian motion on spheres

This section and the following one are devoted to the diffusion equation

\[
(16) \quad \frac{du}{dt} = \frac{1}{(1-x)^{\alpha}(1+x)^{\beta}} \frac{\partial}{\partial x} \left[ (1-x)^{\alpha \cdot 1}(1+x)^{\beta \cdot 1} \frac{\partial u}{\partial x} \right]
\]
on the interval $-1 < x < 1$, with the boundary conditions

$$(1-x)^{\alpha+1} \frac{\partial u}{\partial x} \bigg|_{x=1} = 0, \quad (1+x)^{\beta+1} \frac{\partial u}{\partial x} \bigg|_{x=-1} = 0.$$ 

In general, we are interested in real $\alpha, \beta > -1$. Of particular interest is the ultraspherical case $\alpha = \beta = \gamma - 1/2$, when (16) becomes

$$(17) \quad \frac{\partial u}{\partial t} = (1-x^2) \frac{\partial^2 u}{\partial x^2} - (2\gamma+1)x \frac{\partial u}{\partial x}.$$ 

The relevant fundamental solution (17) has been studied by Bochner [1]; our notation will conform to his.

If $P_n^{(\alpha, \beta)}(x)$ are the Jacobi polynomials ([2], [16]), we renormalize them so that they all have the value 1 at $x = 1$, and use the symbol $P_n^{(\alpha, \beta)}(x)$ for the polynomials so obtained. When there is no danger of confusion we write simply $P_n(x)$; thus

$$P_n(x) = P_n^{(\alpha, \beta)}(x) = P_n^{(\alpha, \beta)}(x)/P_n^{(\alpha, \beta)}(1).$$

The fundamental solution of (16) for the stated boundary conditions is

$$(18) \quad p(t;x,y) = \sum_{n=0}^{\infty} e^{-\lambda_n t} P_n(x)P_n(y)w_n,$$

where $\lambda_n = n(n+\alpha+\beta+1)$ and

$$(19) \quad w_n = \frac{\Gamma(\beta+1)}{\Gamma(\alpha+1)\Gamma(\alpha+\beta+1)} \cdot \frac{\Gamma(n+\alpha+\beta+1)}{\Gamma(n+1)\Gamma(n+\beta+1)} \cdot \frac{2n+\alpha+\beta+1}{\alpha+\beta+1}.$$
The associated transition probability function is

\[
F(t;x,E) = \int_E p(t;x,y) \rho(y) dy,
\]

where

\[
d\rho(y) = \frac{(1-y)\alpha(1+y)\beta dy}{\int_{-1}^{1} (1-x)\alpha(1+x)\beta dx}.
\]

In the remainder of this section we describe a geometrical construction which relates the transition probability (20) for the special case \( \alpha = \beta = (N-1)/2 \) (\( N = 0,1,2,\ldots \)) with a Brownian motion on the surface of the unit sphere \( \Omega \) in a Euclidean space of \( N + 2 \) dimensions. If \( \xi_1, \ldots, \xi_{N+2} \) are Cartesian coordinates on \( \Omega \) such that \( \xi_1^2 + \cdots + \xi_{N+2}^2 = 1 \), then spherical coordinates \( \theta_1, \ldots, \theta_N, \varphi \) are defined by

\[
\begin{align*}
\xi_1 & = \cos \theta_1, \\
\xi_2 & = \sin \theta_1 \cos \theta_2, \\
\vdots & \\
\xi_N & = \sin \theta_1 \sin \theta_2 \cdots \sin \theta_{N-1} \cos \theta_N, \\
\xi_{N+1} & = \sin \theta_1 \cdots \sin \theta_N \cos \varphi, \\
\xi_{N+2} & = \sin \theta_1 \cdots \sin \theta_N \sin \varphi,
\end{align*}
\]

where \( 0 \leq \theta_i \leq \pi \) and \( 0 \leq \varphi \leq 2\pi \). The Laplace-Beltrami operator \( \Delta \) on \( \Omega \) is given by
\[ \Delta u = (\sin \theta_1)^{-N} \frac{\partial}{\partial \theta_1} \left[ (\sin \theta_1)^N \frac{\partial u}{\partial \theta_1} \right] + (\sin \theta_1)^{-2} (\sin \theta_2)^{1-N} \frac{\partial}{\partial \theta_2} \left[ (\sin \theta_2)^{N-1} \frac{\partial u}{\partial \theta_2} \right] + \cdots + (\sin \theta_1 \cdots \sin \theta_{N-1})^{-2} (\sin \theta_N)^{-1} \frac{\partial}{\partial \theta_N} \left[ \sin \theta_N \frac{\partial u}{\partial \theta_N} \right] + (\sin \theta_1 \cdots \sin \theta_N)^{-2} \frac{\partial^2 u}{\partial \varphi^2} \right] \]

The diffusion equation \( \partial u / \partial t = \Delta u \) has a unique fundamental solution on \( \Omega \), call it \( p(t; \xi, \eta) \), and this solution is the transition probability density of a stationary Markov process on \( \Omega \) (the term "density" means relative to the Lebesgue volume element \( d\omega \) on \( \Omega \)). We take it as known that this process \( \xi(t) \), called Brownian motion on \( \Omega \), has continuous path functions. The transition density has the representation

\[ p(t; \xi, \eta) = \sum_{n=0}^{\infty} \sum_{\ell=0}^{n} \frac{\lambda_n}{\pi_n} \int_{0}^{\infty} S_n^{(\ell)}(\xi) S_n^{(\ell)}(\eta) e^{-\lambda_n t} dt, \]

where \( \lambda_n = n(n+\alpha+\beta+1) = n(n+N) \), \( h(n) = (n+N-1)!/(2n+N)! \), and \( S_n^{(\ell)}(\xi) \) \( \ell = 1,2,\ldots,h(n) \) is a complete orthonormal set of surface harmonics of degree \( n \) [2]. By virtue of the addition theorem for surface harmonics [2, p. 243], we have the alternative representation

\[ p(t; \xi, \eta) = \sum_{n=0}^{\infty} e^{-\lambda_n t} \frac{h(n)}{\omega(N)} P_n \left( (\xi, \eta) \right), \]

where \( (\xi, \eta) = \xi_1 \eta_1 + \cdots + \xi_{N+2} \eta_{N+2} \) is the cosine of the angle between
the unit vectors $\xi$ and $\eta$, and $\omega(N) = \int_{\Omega} d\omega$ is the surface area of the unit sphere in $N + 2$ dimensions.

Now we start a Brownian motion $\xi(t)$ on the sphere with an initial distribution which has the $\xi_1$ axis as an axis of symmetry, and which has no mass at the points $\xi_1 = \pm 1$. The distribution of $\xi(t)$ for any $t > 0$ is then also symmetric about the $\xi_1$ axis. The random variable $X(t) = \xi_1(t)$, which is the projection of $\xi(t)$ on the $\xi_1$ axis, is evidently a Markoff process with continuous path functions and with state space $-1 \leq x \leq 1$. To calculate the transition probability function of $X(t)$ we introduce polar coordinates

$$\xi \sim (\theta_1, \theta_2, \ldots, \theta_N, \varphi), \quad \eta \sim (\theta_1', \theta_2', \ldots, \theta_N', \varphi')$$

according to (21). The transition probability is clearly given by

$$P(t; x, [-1, y]) = \Pr\{X(t) \leq y | X(0) = x\} = \frac{1}{\omega(N)} \int_{-1 \leq \eta_1 \leq y} p(t; \xi, \eta) d\omega \eta,$$

where $\xi$ is any fixed unit vector with $\xi_1 = x$. For symmetry considerations we may set $\theta_1 = \cos^{-1} x$ and $\theta_2 = \theta_3 = \ldots = \theta_N = 0$ without loss of generality; hence

$$(\xi, \eta) = \cos \theta_1 \cos \theta_1' + \sin \theta_1 \sin \theta_1' \cos \theta_2' .$$

Using the area element

$$\omega \eta = (\sin \theta_1')^N (\sin \theta_2')^{N-1} \ldots (\sin \theta_N') d\theta_1' d\theta_2' \ldots d\theta_N' d\varphi'$$

(23)
and the representation (22) together with (23), we obtain

\[ P(t;x,[-1,y]) = \int_{\cos \theta_1 - y}^{\cos \theta_1} (\sin \theta_1)^N d\theta_1 \sum_{n=0}^{\infty} e^{-\lambda_n t} \frac{\omega(N-2)\delta(n)}{\omega(N)} f_n(\theta_1,\theta_1'), \]

where

\[ f_n(\theta_1,\theta_1') = \int_0^\pi P_n(\cos \theta_1 \cos \theta_1' + \sin \theta_1 \sin \theta_1' \cos \theta_2')(\sin \theta_2')^{N-1} d\theta_2', \]

(This result has a slightly different form if \( N = 0 \)). Now by virtue of an identity of Gegenbauer (see [17, p. 369]), we have

\[ f_n(\theta_1,\theta_1') = P_n(\cos \theta_1)P_n(\cos \theta_1') \int_0^\pi (\sin \theta_2')^{N-1} d\theta_2', \]

and it readily follows that

\[ P(t;x,[-1,y]) = c_{N-1}^{-1} \int_{-1}^{y} p(t;x,z)(1-z^2)^{(N-1)/2} dz, \]

with \( p(t;x,z) \) as in (18), where \( \alpha = \beta = (N-1)/2 \) and

\[ c_N = \int_{-1}^{1} (1-t^2)^{N/2} dt. \]

Thus by appealing to the coincidence probability theorem we conclude that the density (18) is totally positive for these special values of \( \alpha, \beta \).

4. **Hahn polynomials**

In the previous section, we proved that (18) is totally positive in the special cases \( \alpha = \beta = (N-1)/2 \), with \( N \) an integer. In order to
prove total positivity for general values of the parameter, we shall
discretize the Jacobi diffusion process defined by (16). This leads to
a related birth and death process of independent interest [10]. The
polynomials associated with this process are the Hahn polynomials, finite
systems of orthogonal polynomials that approximate the Jacobi polynomials.
With fixed real \( \alpha, \beta > -1 \) and integer \( N \geq 1 \), let
\[
\rho_N(x) = \frac{\binom{\alpha+x}{x} \binom{\beta+N-1-x}{N-1}}{\binom{N+\alpha+\beta}{N-1}}.
\]
Then the Hahn polynomials \( Q_n(x) = Q_n(x; \alpha, \beta, N) \) \( (n = 0, 1, \ldots, N-1) \) are
determined to within constant factors of modulus 1 by the orthogonality relation
\[
\sum_{x=0}^{N-1} Q_n(x)Q_m(x)\rho_N(x) = \frac{\delta_{m,n}}{w_{N,n}},
\]
where
\[
w_{N,n} = w_n^{\binom{N-1}{n}/\binom{N+\alpha+\beta+n}{n}}
\]
and \( w_n \) is as in (19). We normalize the polynomials by the condition
\( Q_n(0) = 1 \) \( (n = 0, 1, \ldots, N-1) \). This system of polynomials has been
investigated by several authors (see [2] and the references given there),
yet several of the properties needed below appear to be new. We next
state several formulas without proof (we shall publish elsewhere a more
extensive account of the properties of the Hahn polynomials). The
explicit formula
\[ Q_n(x) = \binom{\alpha+\beta+n+1}{\alpha+1} \binom{\alpha+\beta+n+1}{\beta+n+1} \]

was given by Weber and Erdélyi [18]. Of central interest here is the difference equation

\[ -\lambda_n Q_n(x) = D(x)Q_n(x-1) - [B(x) + D(x)]Q_n(x) + B(x)Q_n(x+1), \]

with \( B(x) = (\alpha+1+x)(N-1-x) \), \( D(x) = x(N+\beta-x) \), and \( \lambda_n = n(n+\alpha+\beta+1) \).

Only the case \( \alpha = \beta = 0 \) of (25) is cited in [2, p. 223]. This equation, which is satisfied for all complex \( x \), may be written in the form

\[ -\lambda_n Q_n(x) = \frac{1}{\rho_N(x)} \Delta^- \{ B(x)\rho_N(x)\Delta^+[Q_n(x)] \}, \]

where \( \Delta^- \) and \( \Delta^+ \) are the difference operators defined by

\[ \Delta^+[f(x)] = f(x+1) - f(x), \quad \Delta^-[f(x)] = f(x) - f(x-1). \]

The limit formula

\[ \lim_{N \to \infty} Q_n((N-1)x; \alpha, \beta, N) = P_n(1-2x; \alpha, \beta) \]

connects the Hahn polynomials with the Jacobi polynomials, the convergence being uniform in every bounded region of the complex \( x \) plane for fixed \( n \). This implies

\[ \lim_{N \to \infty} (N-1)[Q_n((N-1)x+1) - Q_n((N-1)x)] = \frac{d}{dx} P_n(1-2x) \]
and similar relations for the higher derivatives. Since for $0 < x < 1$

$$N\rho_N((N-1)x) \sim \frac{x^{\alpha(1-x)\beta}}{\int_0^1 x^{\alpha(1-x)\beta}dx}$$

we find that the difference equation of the Hahn system passes over formally into the differential equation of the Jacobi system when $N \to \infty$.

Now the Hahn polynomials can be associated with a finite birth and death process on the states $0, 1, \ldots, N-1$ which has as its backward equation

$$\frac{\partial u_k}{\partial t} = D(k)u_{k-1} - [B(k)+D(k)]u_k + B(k)u_{k+1} \quad (k = 0, \ldots, N-1),$$

or, more compactly, $\frac{\partial u}{\partial t} = Au$, where $u$ is a column vector with $N$ entries and $A$ is the obvious $n$-square Jacobi matrix of birth and death type. The transition probability matrix $P(t) = (P_{ij}(t))$ of the birth and death process is the solution of $dP/dt = AP$, subject to the initial condition $P(0) = I$, and for it we have the two alternative representations $P(t) = e^{tA}$ and

$$P_{ij}(t) = \rho_N(j) \sum_{n=0}^{N-1} \frac{-\lambda^t}{n!} Q_n(i)Q_n(j)w_{N,n}.$$  

The matrix $P(t)$ is known to be totally positive [7; see Lemma 10, p. 541]. We shall show that the density (18) may be obtained as a limit from the "density"

$$p(t;i,j;N) = \frac{P_{ij}(t)}{\rho_N(j)} = \sum_{n=0}^{N-1} \frac{-\lambda^t}{n!} Q_n(i)Q_n(j)w_{N,n}.$$
From (24) follows the crude bound
\[ |Q_n(x)| \leq \left( 2 + \frac{\beta + n}{2 + 1} \right)^{2n} (x = 0, 1, \ldots, N-1) . \]

Moreover, \( 0 < w_{N,n} < w_{N+1,n} \to w_n \) as \( N \to \infty \), and \( w_n = o(n^{2\alpha + 1}) \) for \( n > 1 \). These estimates enable us to conclude that the series
\[ \sum_{n=0}^{N-1} e^{-\lambda_n t} Q_n(x) Q_n(y) w_{N,n} (x, y = 0, 1, \ldots, N-1 ; N \geq 1) \]
are termwise dominated by the convergent series
\[ \sum_{n=0}^{\infty} Ce^{-\lambda_n t} \left| 2 + \frac{\beta + n}{2 + 1} \right|^n n^{2\alpha + 1} \]
for some constant \( C \). Now if \( x \) and \( y \) are rational numbers in \( [0, 1] \), then we can choose a sequence of integers \( \{N_r\} \) such that \( (N_r - 1)x \) and \( (N_r - 1)y \) are integers for each \( r \) and \( N_r \to \infty \) as \( r \to \infty \). We form
\[ p(t; (N_r - 1)x, (N_r - 1)y; N_r) = \sum_{n=0}^{N_r} e^{-\lambda_n t} Q_n((N_r - 1)x) Q_n((N_r - 1)y) w_{N_r,n} . \]

By virtue of (26) and the above domination, we have
\[ \lim_{r \to \infty} p(t; (N_r - 1)x, (N_r - 1)y; N_r) = \sum_{n=0}^{\infty} e^{-\lambda_n t} P_n(1-2x) P_n(1-2y) w_n \]
\[ = p(t; 1-2x, 1-2y) . \]

Thus (18) is non-negative for rational \( x, y \) in \( [0, 1] \) and by continuity for all \( x, y \) in the interval. Positivity of higher order follows easily in a similar manner.
Finally we show that (18) is strictly totally positive in the sense that if \(-1 \leq x_1 < \ldots < x_m \leq 1\) and \(-1 \leq y_1 < \ldots < y_m \leq 1\), then

\[
p(t; x_1, \ldots, x_m \ y_1, \ldots, y_m) > 0 .
\]

Consider first the case \(m = 1\). The function \(p(t; x, y)\) is continuous on the square \(-1 \leq x, y \leq 1\), and along the diagonal \(x = y\) it is everywhere positive; in fact, it follows from (18) that

\[
p(t; x, x) \geq e^{-\lambda_0 t} P_0(x)P_0(x)w_0 = 1 .
\]

Hence there is a neighborhood of the diagonal, say

\[
U = \{(x, y); \ x-y| < \epsilon\} ,
\]

in which \(p(t; x, y)\) is positive. For any point \((x, y)\) in the square there is a finite set \(y_1, \ldots, y_k\) such that all the points \((x, y_1), (y_1, y_2), \ldots, (y_{k-1}, y_k), (y_k, y)\) lie in \(U\). The Chapman-Kolmogorov equation gives

\[
p(t; x, y) = \int_{-1}^{1} \cdots \int_{-1}^{1} p\left(\frac{t}{k+1}; x, y_1\right) p\left(\frac{t}{k+1}; y_1, y_2\right) \cdots p\left(\frac{t}{k+1}; y_{k-1}, y_k\right) \rho(y_1) \cdots \rho(y_k) dy_1 \cdots dy_k ,
\]

from which it can be seen that \(p(t; x, y) > 0\). To deal with the case \(m > 1\), we introduce the determinental polynomials
\[ P(x_1, \ldots, x_m) = \det P_{n_1}(x_j) \ (0 \leq n_1 < n_2 < \ldots < n_m) \]

In terms of these polynomials there is a representation

\[ p(t; y_1, \ldots, y_m) = \sum_{0 < n_1 < \ldots < n_m} \exp \left\{ -(\lambda_{n_1} + \ldots + \lambda_{n_m}) t \right\} P(x_1, \ldots, x_m) \]

\[ \left[ \begin{array}{c} n_1 \, n_m \end{array} \right] \]

\[ \prod_{y_i, y_j} \]

\[ c \prod_{i < j} (x_j - x_i) \ (c \neq 0) , \]

from which it follows that

\[ p(t; x_1, \ldots, x_m) > 0 \ (-1 \leq x_1 < \ldots < x_m \leq 1) . \]

There is a Chapman-Kolmogorov equation of the form

\[ p(t+s; y_1, \ldots, y_m) = \int \ldots \int p(t; z_1, \ldots, z_m) p(s; y_1, \ldots, y_m) \rho(z_1) \ldots \rho(z_m) dz_1 \ldots dz_m \]

and with this equation, using the case \( m = 1 \) as a model, it is easy to construct an argument verifying (27).

Wright [19] and Kimura [11] have proposed various stochastic models of chromosome division. One of these models has as the state variable the proportion of the number of mutant subunits in a chromosome cell. Specifically, we consider a model in which each chromosome consists of \( n \)
subunits and suppose that a mutation has occurred in one of them. The subunits duplicate to produce \( 2n \), which divide at random into two daughter chromosomes of \( n \) subunits. We trace a single line of descent. The state \( E_{ij} \) \( (i = 0, 1, \ldots, n) \) in each generation designates the number of mutant subunits contained in the cell. The transition probabilities are given by

\[
P_{ij} = \frac{\binom{2n}{i} \binom{2n-2i}{n-i}}{\binom{2n}{n}}.
\]

If \( n \) is very large, the proportion of mutant subunits \( x = i/n \), \( 0 \leq x \leq 1 \) can be regarded as a continuous variable. Let \( \phi(x, t) \) be the probability density of \( x \) at time \( t \). If \( \delta x \) is the amount of change in \( x \) per generation, then

\[
E(\delta x) = 0, \quad E((\delta x)^2) = x(1-x)/(2n-1)
\]

and \( E((\delta x)^k) = o(1/n) \) for \( k \geq 3 \). A continuous approximation used to obtain \( \phi(x, t) \) is furnished by the forward diffusion equation

\[
(28) \quad \frac{\partial \phi(x, t)}{\partial t} = \frac{1}{2(2n-1)} \frac{\partial^2}{\partial x^2} [x(1-x)\phi(x, t)]
\]

where \( \phi(x, t) \) has the initial distribution function \( \delta \left(x - \frac{1}{2n}\right) \) and \( \delta \) is the Dirac function (see [11]). Clearly (28), except for a translation, is essentially (16) with \( \alpha = \beta = 0 \). Other examples of the same kind which lead to the Jacobi diffusion process are discussed in [3] and [1].
REFERENCES


