Properties of kernels of integral equations whose iterates satisfy linear relations

Carl Eric Langenhop
Iowa State College

Follow this and additional works at: https://lib.dr.iastate.edu/rtd
Part of the Mathematics Commons

Recommended Citation
Langenhop, Carl Eric, "Properties of kernels of integral equations whose iterates satisfy linear relations" (1948). Retrospective Theses and Dissertations. 12897.
https://lib.dr.iastate.edu/rtd/12897
NOTE TO USERS

This reproduction is the best copy available.

UMI®
PROPERTIES OF KERNELS OF INTEGRAL EQUATIONS
WHOSE ITERATES SATISFY LINEAR RELATIONS

by

Carl Eric Langenhop

A Thesis Submitted to the Graduate Faculty
for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: Mathematics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State College
1948
INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.
# TABLE OF CONTENTS

I. INTRODUCTION AND GENERAL THEORY ........................................ 1

II. IDEMPOTENT KERNELS ............................................................ 13

III. THE IDEMPOTENT KERNEL AS A MARKOFF PROCESS .................. 20

IV. GENERAL LINEAR RECURRENCE RELATIONS
    BETWEEN THE ITERATED KERNELS ......................................... 36

V. SUMMARY ................................................................................. 50

VI. ACKNOWLEDGEMENT .............................................................. 51
I. INTRODUCTION AND GENERAL THEORY

An equation of the type

\[ u(x) = f(x) + \lambda \int_{a}^{b} K(x,t)u(t)dt \]  \hspace{1cm} (1)

(the function \(f(x)\) and the "kernel" \(K(x,t)\) are supposed known functions and \(\lambda\) a constant) is known as a linear integral equation of the Fredholm type, after the man who first studied them extensively.

The special case of (1) with \(f(x) = 0\)

\[ u(x) = \lambda \int_{a}^{b} K(x,t)u(t)dt \]  \hspace{1cm} (2)

is called a homogeneous integral equation.

In the classical theory of these equations the function \(K(x,t)\) is usually taken to be continuous. Then if \(f(x)\) is likewise continuous there are theorems to the effect that "there exists one and only one continuous solution" given by a certain formula \(^1\).

Continuity of \(K(x,t)\) is usually assumed for two reasons;

\(^1\) We are not so much concerned here with the solutions of equations (1) and (2) so the explicit form of the solution is not given.
(a) if \( g(t) \) is continuous, then \( \int_a^b K(x,t)g(t)dt \) will also be continuous,

(b) interchanges in the order of integrations are permitted, such as, for example,

\[
\int_a^b \int_a^b K(x,s)K(s,t)dsdt = \int_a^b \int_a^b K(x,s)K(s,t)dtds.
\]

Throughout this paper we are mainly concerned with the kernels \( K(x,t) \) and certain integration operations performed on them. Thus for the sake of generality (with a view to wider application than simply to integral equations) we will assume only certain measurability properties on \( K(x,y) \) and that integration is in the Lebesgue sense. For the purposes of this paper then we make the following

**DEFINITION:** A kernel \( K(x,y) \) is a real function of the two real variables \( x \) and \( y \), which satisfies the following conditions over the square \( a \leq x, y \leq b \):

(a) \( K(x,y) \) is a measurable function of \( (x,y) \),

(b) \( \int_a^b K(t,t)dt \) exists.

The first condition is sufficient (by Fubini's
theorem\(^2\) to allow interchanges in the order of integrations in multiple integrals involving \(K(x,y)\). Furthermore if \(g(t)\) is measurable, then (again a special case of Fubini's theorem)
\[
h(x) = \int_a^b K(x,t)g(t)\,dt
\]
is defined for almost all \(x\) in the interval \((a,b)\) and is an integrable function of \(x\) on this interval.

The second condition on \(K(x,y)\) is needed later so that certain functionals of a kernel will always exist.

With these restrictions on \(K(x,y)\) the theorems quoted in this chapter are still true in the sense of Lebesgue integrations and measurable functions, although the references given establish them only for continuous \(K(x,y)\).

Part of the importance of the linear Fredholm integral equation stems from the fact that many

---

boundary value problems (in particular, ordinary
linear differential equations with certain general
linear boundary conditions) can be expressed in the
more compact form of such integral equations 3.

The literature concerning the above equations
is quite extensive and many facts about their
solutions are known. Most of the known facts clearly
point out the analogy between equations (1) and (2)
and systems of linear algebraic equations. In fact
Fredholm 4 first used this analogy in studying these
integral equations. In his theory the integral
equation is replaced by a system of linear algebraic
equations, the solution of the algebraic system is
found, and the limit of this solution as the system
becomes larger and larger is obtained. The limit
(obtained heuristically) of the solution of the
algebraic system is then shown to be the only
solution of the integral equation.

3 Margenau, H. and Murphy, G.M. "The Mathematics
of Physics and Chemistry". D. VanNostrand Co.,

4 Fredholm. "Sur une classe d'equations fonction-
Let the interval \((ab)\) be divided into \(n\) equal sub-intervals by inserting the points

\[ x_i = a + ih, \quad i = 0, 1, \ldots, n, \quad h = \frac{b - a}{n}. \]

The equation (1) is then replaced by the system of equations

\[ (3) \quad u(t_i) = f(t_i) + \lambda \sum_{j=1}^{n} K(t_i, t_j) u(t_j) h, \quad i = 1, 2, \ldots, n, \]

which can be conveniently expressed in matrix notation as

\[ (3') \quad (I - \lambda hK)U = F \]

where \(K\) is the matrix of elements \(K_{ij} = K(t_i, t_j)\) and \(U\) and \(F\) are the column vectors \((u(t_i))\) and \((f(t_i))\) respectively.

The system \((3')\) can be solved for \(U\) if the determinant

\[ \Delta = |I - \lambda hK| \neq 0. \]

In the solution of the integral equation (1) the limit of the above determinant as \(n \to \infty\) plays an analogous role. This limit, which is, of course, a function of the parameter \(\lambda\), is called the Fredholm determinant of the kernel \(K(x, y)\). It is shown in
standard works on the subject that

\[ (4) \lim \Delta = D(\lambda) = 1 + \sum_{r=1}^{\infty} \frac{(-\lambda)^r}{r!} \int_a^b \cdots \int_a^b K(t_1, \ldots, t_r) dt_1 \cdots dt_r \]

where

\[ K(t_1, t_2, \ldots, t_r) = \begin{vmatrix} K(t_1, s_1) & K(t_1, s_2) & \cdots & K(t_1, s_r) \\ K(t_2, s_1) & K(t_2, s_2) & \cdots & K(t_2, s_r) \\ \vdots & \vdots & \ddots & \vdots \\ K(t_r, s_1) & K(t_r, s_2) & \cdots & K(t_r, s_r) \end{vmatrix} \]

An important property of \( D(\lambda) \) will be used later.

We mention it as

**Theorem 1.1:** If \( K(x, y) \) is a bounded kernel, then

\( D(\lambda) \) is an entire function.

---


Of course, in these derivations integrations are in the sense of Riemann, but they need not be. In the more general sense it is to be understood that \( D(\lambda) \) must be defined only as the power series given in equation (4). All theorems as to existence of solutions etc. in the Fredholm theory are based only on these and other closely related series and not on any limiting quantities such as \( \lim \Delta \). The treatment of the integral equation as a limiting case of a linear algebraic system is, however, the historical approach.
The proof of this makes use of Hadamard's theorem on determinants. It can be found in most books on integral equations ⁶.

Also in integral equation theory is the Fredholm minor \( D(x,y; \lambda) \) analogous to the first minors of the determinant \( \Delta \). The Fredholm minor is defined as

\[
D(x,y; \lambda) = \lambda K(x,y) + \\
\sum_{r=1}^{\infty} \frac{(-1)^r \lambda^{r+1}}{r!} \int_a^b \cdots \int_a^b K(x,t_1,\ldots,t_r) dt_1 \cdots dt_r
\]

About \( D(x,y; \lambda) \) it is known ⁶

**Theorem 1.2:** If \( K(x,y) \) is a bounded kernel, then the series defining \( D(x,y; \lambda) \) converges for all \( \lambda \) and is uniformly convergent in \( x \) and \( y \) in the square \( a \leq x,y \leq b \).

The following two relations between \( D(\lambda) \) and \( D(x,y; \lambda) \) are called Fredholm's two fundamental relations ⁸

---

⁶ Bôcher, op.cit., p. 39.
⁷ Kowalewski, op.cit., p. 137.
Lovitt, op.cit., p. 33.
Bôcher, op.cit., p. 27.
⁸ Bôcher, op.cit., p. 39.
\[(6a) \quad D(x,y; \lambda) - \lambda K(x,y) D(\lambda) = \lambda \int_{a}^{b} D(x,t; \lambda) K(t,y) dt\]

\[(6b) \quad = \int_{a}^{b} K(x,t) D(t,y; \lambda) dt,\]

These will be needed later.

In the homogeneous case of the integral equation (i.e. equation (2)) the linear algebraic system (3') becomes

\[(I - \lambda K) U = 0.\]

As is well known, in order for this to have a non-trivial solution (not all \(u_i = 0\)) we must have

\[\Delta = |I - \lambda K| = 0.\]

Generally this will be true only for certain values of \(\lambda\).

Similarly it has been shown \(^9\) that equation (2) has a non-trivial solution \(u(x) \neq 0\) only for those values of \(\lambda\) for which \(D(\lambda) = 0\). These values are called the characteristic constants of the kernel \(K(x,y)\), and the solutions of (2) when \(\lambda = \lambda_\rho\) (\(\lambda_\rho\) a characteristic constant) are called the characteristic functions of \(K(x,y)\) belonging to \(\lambda_\rho\). For a continuous \(K(x,y)\), in general if \(\lambda_\rho\) is a root of \(D(\lambda)\) of

\(^9\) Ibid., p. 44.
multiplicity \( k_p \), there are \( k_p \) linearly independent continuous solutions of (2) belonging to \( \lambda_p \).

In some phases of the theory the iterated kernels play a large part. These are defined by the relations:
\[ K_1(x,y) = K(x,y), \]
\[ K_n(x,y) = \int_a^b K_{n-1}(x,t)K(t,y)dt = \int_a^b K(x,t)K_{n-1}(t,y)dt, \ n \geq 1. \]

It is interesting to note that the coefficients of the various powers of \( \lambda \) in \( D(\lambda) \) and \( D(x,y; \lambda) \) can be expressed as relatively simple linear combinations of the iterated kernels or their traces. Since we will need these expressions we develop them now.

If for convenience we define
\[ A_0 = 1, \]
\[ A_n = \int_a^b \cdots \int_a^b K(t_1, \ldots, t_n)dt_1 \cdots dt_n, \ n \geq 1. \]

10 Lovitt, op. cit., pp. 46-55.

11 This equation holds by virtue of the conditions put on \( K(x,y) \). Thus, for instance,
\[ \int_a^b K_n(x,t)K(t,y)dt = \int_a^b (\int_a^b K(x,s)K(s,t)ds)K(t,y)dt \]
\[ = \int_a^b K(x,s) (\int_a^b K(s,t)K(t,y)dt)ds \]
\[ = \int_a^b K(x,s)K_n(s,y)ds. \]
and
\[
B_n(x,y) = \int_{a}^{b} \cdots \int_{a}^{b} K\left( x, t_1, \ldots, t_n \right) \, dt_1 \cdots dt_n, \quad n \geq 0,
\]
then \( D(\lambda) \) and \( D(x,y;\lambda) \) can be written
\[
D(\lambda) = \sum_{n=0}^{\infty} \frac{(-\lambda)^n}{n!} A_n,
\]
\[
D(x,y;\lambda) = \sum_{n=0}^{\infty} \frac{(-1)^n \lambda^n \lambda^{n+1}}{n!} B_n(x,y).
\]

Furthermore we see from the definition of \( A_n \) and \( B_n(x,y) \) that
\[
(7) \quad A_{n+1} = \int_{a}^{b} B_n(x,x) \, dx, \quad n \geq 0.
\]

Now using the second of Fredholm's fundamental relations, viz. equation (6b) and equating coefficients of like powers of \( \lambda \) we obtain\(^{12}\)
\[
(8) \quad B_n(x,y) = A_n K(x,y) - n \int_{a}^{b} K(x,t) B_{n-1}(t,y) \, dt, \quad n \geq 1.
\]

This enables us to deduce

**THEOREM 1.3:** The following relation holds between the \( A_n \), \( B_n(x,y) \) and \( K_n(x,y) \);
\[
(9) \quad B_n(x,y) = A_n K_n(x,y) - nA_{n-1} K_n(x,y) + n(n-1)A_{n-2} K_n(x,y) + \cdots + (-1)^n n! K_{n+1}(x,y).
\]

\(^{12}\) Actually Fredholm's fundamental relations are obtained from equation (8) and a similar one, both of which follow rather easily from the definitions of the \( A_n \) and the \( B_n(x,y) \) and the properties of determinants.
Proof: The relation (9) is certainly true for \( n = 1 \) since

\[
B_1(x,y) = \int_a^b \frac{K(x,y)}{K(t,y)} K(t,t) \, dt
\]

\[
= \int_a^b [K(x,y)K(t,t) - K(x,t)K(t,y)] \, dt
\]

\[
= A_1K_1(x,y) - K_2(x,y).
\]

For convenience we write (9) in the form

\[
B_n(x,y) = \sum_{r=0}^{n} \frac{(-1)^r n!}{(n-r)!} A_{n-r}K_{r+1}(x,y).
\]

Now assume that this holds for \( n \). Then by (8)

\[
B_{n+1}(x,y) = A_{n+1}K_1(x,y)
\]

\[
- (n+1) \int_a^b \frac{K(x,t)}{K(t,y)} \sum_{r=0}^{n} \frac{(-1)^r n!}{(n-r)!} A_{n-r}K_{r+1}(t,y) \, dt
\]

\[
= A_{n+1}K_1(x,y) + \sum_{r=0}^{n} \frac{(-1)^{r+1}(n+1)!}{(n-r)!} A_{n-r}K_{r+2}(x,y)
\]

\[
= \sum_{r=0}^{n+1} \frac{(-1)^r (n+1)!}{(n+1-r)!} A_{n+1-r}K_{r+1}(x,y).
\]

Thus the relation then also holds for \( n + 1 \). By induction, therefore, (9) is established for all integral \( n \geq 1 \).

COROLLARY 1.3.1: The following relation holds between the \( A_n \):

\[
\text{The following relation holds}
\]
(10) \[ A_n = A_{n-1}P_1 - (n-1)A_{n-2}P_2 + \ldots + (-1)^{n-1}(n-1)! P_n, \]

where

\[ P_r = \int_a^b K_r(t,t)dt. \]

**Proof:** This follows immediately from Theorem 1.3 and equation (7) since

\[ A_n = \int_a^b B_{n-1}(x,x)dx. \]

A special class of kernels are those of the form

\[ K(x,y) = \sum_{i=1}^N a_i(x)\beta_i(y), \]

which we shall designate as a separable kernel\(^{14}\).

For such kernels the solving of the integral equation (1) and (2) can be reduced exactly to

For a kernel the existence of \( P_1 = \int_a^b K(t,t)dt \) was assumed, but the existence of the other \( P_r, r \geq 2 \), follows from the assumption of measurability of \( K(x,y) \). Thus for instance, there exists

\[ \int_a^b \int_a^b K(x,t)K(t,x)dxdt = \int_a^b \left( \int_a^b K(x,t)K(t,x)dt \right) dx \]

\[ = \int_a^b K_a(x,x)dx = P_a, \]

and similarly for the other \( P_r \).

\(^{14}\) Kowalewski calls such kernels "kernels with separated variables".
the solving of a finite system of linear algebraic equations \(^{15}\) (though not in the same manner as the system (3) was obtained). In fact it is known that for a separable kernel \(D(\lambda)\) is a polynomial, i.e. it has only a finite number of terms. \(^{16}\)

Since there are non-separable kernels for which \(D(\lambda)\) is a polynomial, it would be of interest to know some additional properties of \(D(\lambda)\) and \(D(x,y;\lambda)\) which are true only if \(K(x,y)\) is separable. One of the results of this investigation is an answer to this question.

\(^{15}\) Kowalewski, op. cit., pp. 142-43.

\(^{16}\) Ibid., pp. 143-51.
II. IDEMPOTENT KERNELS

We define an idempotent kernel to be one which satisfies the relation

\[ \int_{a}^{b} K(x,t)K(t,y)dt = K(x,y), \]

for \( x \) and \( y \) in the square \( a \leq x, y \leq b \).

This equation in itself might be considered as a non-linear integral equation. Also here the function \( K(x,y) \) is both the kernel and the solution of a linear integral equation, and by considering it as a kernel we can determine some of its properties as a solution and vice-versa. This we proceed to do by proving.

**Theorem 2.1:** For any idempotent kernel \( K(x,y) \)

\[ D(\lambda) = (1 - \lambda)A_1 \]

where

\[ A_1 = \int_{a}^{b} K(t,t)dt. \]

**Proof:** We notice that for a kernel satisfying (11)

\[ K_a(x,y) = \int_{a}^{b} K(x,t)K(t,y)dt = K(x,y) \]

and in general

\[ K_a(x,y) = K(x,y). \]
We have then
\[ p_a = \int_{a}^{b} K_n(t, t) dt = \int_{a}^{b} K(t, t) dt = A_1, \]
so for this kernel equation (10) reads

(12) \[ A_n = A_{n-1} A_1 - (n-1) A_{n-2} A_1 + \cdots + (-1)^{n-1}(n-1)! A_1 \]

Also

(13) \[ A_{n+1} = A_1 (A_n - n A_{n-1} + n(n+1) A_{n-2} + \cdots + (-1)^n n!). \]

Multiplying equation (12) by \( n \) and adding this to equation (13) we get

\[ A_{n+1} + n A_n = A_1 A_n \]

or

(14) \[ A_{n+1} = A_n (A_1 - n). \]

We suppose then that

(15) \[ A_n = A_1 (A_1 - 1) \cdots (A_1 - n + 1), \]

which obviously holds for \( n = 1 \). By (14) and (15)

\[ A_{n+1} = A_n (A_1 - n) \]

\[ = A_1 (A_1 - 1) \cdots (A_1 - n + 1)(A_1 - n + 1 + 1) \]

so that (15) then also holds for \( n + 1 \). By mathematical induction equation (15) is established for all integral \( n \geq 1 \).

Thus by the definition of \( D(\lambda) \) (equation (4); also see page 9) we have that for an idempotent kernel
\[ D(\lambda) = 1 - A_1 \lambda + \frac{A_1(A_1-1)}{2!} \lambda^2 - \ldots + (-1)^n \frac{A_1(A_1-1) \cdots (A_1-n+1)}{n!} \lambda^n + \ldots \]

\[ = (1-\lambda)^{A_1}. \]

**COROLLARY 2.1.1:** If \( K(x,y) \) is bounded and is idempotent, then \( \int_a^b K(t,t)dt \) is a positive integer or zero.

**Proof:** This follows immediately since by Theorem 1.1 \( D(\lambda) \) is an entire function when \( K(x,y) \) is bounded. It is well known that \( (1-\lambda)^{A_1} \) is an entire function of \( \lambda \) if and only if \( A_1 \) is a positive integer or zero.

It should perhaps be remarked here that for an idempotent kernel \( D(\lambda) \) can also be obtained in a somewhat simpler way. We have already noticed that for such a kernel \( P_n = A_1 \). Now for small enough \( \lambda \) it is known that \(^{17}\)

\[ D(\lambda) = e^{-P_1 \lambda - \frac{1}{2} P_2 \lambda^2 - \ldots - \frac{1}{n!} P_n \lambda^n} \]

which for an idempotent kernel becomes

---

\(^{17}\) Kowalewski, op.cit., p. 124.
This, however, must be $D(\lambda)$ for all $\lambda$ since $D(\lambda)$, being entire, must also be analytic in the whole complex plane. The only analytic function represented by $(1-\lambda)^{A_1}$ for small $\lambda$ is this same function with $A_1$ a positive integer or zero.

From the form of $D(\lambda)$ we see that $\lambda = 1$ is the only characteristic constant of an idempotent kernel. Moreover, the multiplicity of this root being $A_1$ suggests that there are at most $A_1$ linearly independent solutions of (2) for such kernels. From equation (11), however, we notice that for any fixed $y = y_1$ say, $u(x) = K(x, y_1)$ is indeed a solution of (2). This enables us to prove

**Theorem 2.2:** Every bounded idempotent kernel is separable.

**Proof:** As mentioned before, for a fixed $y_1$ on the interval $(a,b)$, $u(x) = K(x, y_1)$ is a solution of

\[
\int_{a}^{b} K(x, t)u(t)dt = u(x), \quad a \leq x \leq b.
\]
If \( u(x) = 0 \) for almost all \( x \), then from this equation we see that \( u(x) = 0 \). Thus unless \( u(x) \) is a trivial solution of (16) it is different from zero on a set of positive measure, and we have

\[
\int_{a}^{b} u^2(t) dt > 0.
\]

For different \( y_\alpha \), \( K(x,y_\alpha) \), are solutions of (16) as is any linear combination of them. Furthermore any non-trivial linear combination of the \( K(x,y_\alpha) \) would have the property (17). Suppose now that there are an infinite number of linearly independent \( K(x,y_\alpha) \) (considered as functions of \( x \)). There certainly exists then as large a finite number of such \( K(x,y_\alpha) \) as we may specify. If \( |K(x,y)| < k \) consider any set of linearly independent non-trivial \( K(x,y_1) \), \( i = 1,2,\ldots n \) where \( n > k^2(b-a)^2 \). There exists then linear combinations of these \( K(x,y_1) \) forming an orthonormal series of solutions of (16), \( u_1(x), u_2(x), \ldots u_n(x) \); i.e.

functions such that

\[
\int_{a}^{b} u_i(t)u_j(t) dt = \delta_{ij},
\]

each \( u_i(x) \) satisfying equation (16).

Now by Bessel's inequality we have

\[
\sum_{i=1}^{n} [u_i(x)]^2 = \sum_{i=1}^{n} \left[ \int_{a}^{b} K(x,t)u_i(t)dt \right]^2 \leq \int_{a}^{b} [K(x,t)]^2 dt
\]

since the \( u_i(x) \) are orthonormal. Hence

\[
n = \sum_{i=1}^{n} \int_{a}^{b} [u_i(x)]^2 dx \leq \int_{a}^{b} \int_{a}^{b} [K(x,t)]^2 dtdx \leq k^2 (b-a)^2.
\]

This contradicts the original choice of \( n \) so we must conclude that there are only a finite number of linearly independent \( K(x,y) \).

Let \( N \) be the number of linearly independent \( K(x,y) \) and let \( u_i(x), i = 1, 2, \ldots, N \) be an orthonormal system of solutions constructed from them. For each \( y \) on the interval \((a,b)\), \( K(x,y) \) must then be a linear combination of the \( u_i(x) \); that is

\[
K(x,y) = \sum_{i=1}^{N} u_i(x)v_i(y).
\]

Thus we have shown \( K(x,y) \) to be separable.

From the expansion (18) we can deduce even more about \( K(x,y) \). We have in particular

\[
\int_{a}^{b} K(x,y)u_1(x)dx = \sum_{j=1}^{N} v_j(y) \int_{a}^{b} u_1(x)u_j(x)dx = v_1(y).
\]

Also
\[
\int_a^b v_1(t)u_j(t)dt = \int_a^b u_j(t) \int_a^b K(x,t)u_1(x)dxdt
= \int_a^b u_1(x) \int_a^b K(x,t)u_j(t)dtdx
= \int_a^b u_1(x)u_j(x)dx = \delta_{ij}.
\]

These results we state as

**Theorem 2.3:** Every bounded idempotent kernel can be written as

\[
K(x,y) = \sum_{i=1}^N u_i(x)v_i(y)
\]

with

\[
\int_a^b u_1(t)u_j(t)dt = \int_a^b u_1(t)v_j(t)dt = \delta_{ij}.
\]

**Corollary 2.3.1:** For a bounded idempotent kernel, \(\int_a^b K(t,t)dt\) is a positive integer or zero.

**Proof:** From Theorem 2.3

\[
\int_a^b K(t,t)dt = \sum_{i=1}^N \int_a^b u_i(t)v_1(t)dt = N.
\]
III. THE IDEMPOTENT KERNEL AS A MARKOFF PROCESS

The idempotent Markoff stochastic process in the continuous case can under certain circumstances be considered as an instance of a function or kernel whose iterates satisfy a linear relation. Such a process as defined by Blackwell is one whose defining probability measure satisfies the relation

\[ (19) \quad \int_{S} P(t,E)P(x,dt) = P(x,E). \]

For any \( x \) in the space \( S \), \( P(x,E) \) is a probability measure over \( S \), that is \( P(x,E) \) is a completely additive set function over the class of all Borel sets in \( S \) and

\[ (19') \quad 0 \leq P(x,E) < 1 \quad \text{with} \quad P(x,S) = 1. \]

If a probability density function exists, i.e., a function \( p(x,y) \) such that

\[ (20) \quad P(x,E) = \int_{E} p(x,y)dy \]

---

then conditions (19) and (19') can be expressed

in the equivalent form

\[(21) \quad \int_S p(x,t)p(t,y)dt = p(x,y)\]

with

\[(21') \quad p(x,t) \geq 0 \quad \text{and} \quad \int_S p(x,t)dt = 1.\]

In this case Blackwell has proved the following

THEOREM 3.1: If \( p(x,y) \) is an idempotent density

function (conditions (3) and (3')), then

\[ S = F + V + A_1 + A_2 + \ldots \]

where

a) \( p(x,y) = p_a(y) \quad \text{for} \ x \ \text{in} \ A_a \)
b) \( p_a(y) > 0 \quad \text{for} \ y \ \text{in} \ A_a \)
c) \( \int_{A_n} p_a(y)dy = 1 \)
d) \( p(x,y) = 0 \quad \text{for} \ y \ \text{in} \ F \)
e) \( m(V) = 0. \)

Doob\(^2^2\) gives a proof of the analogous theorem

\(^{20}\) Ibid., p. 565.

\(^{21}\) Ibid., p. 565-7.

for a matrix \( U = (u_{ij}) \) satisfying the conditions

\[
\sum_{j} u_{ij} \leq 1, \quad u_{ij} \geq 0.
\]

We give here an alternate proof (somewhat analogous to Doob's) of Blackwell's result in the case of a bounded idempotent kernel \( K(x,y) \). Thus we assume

\[
\begin{align*}
\int_{a}^{b} K(x,t)K(t,y)\,dt &= K(x,y), \quad a \leq x, y \leq b, \\
\int_{a}^{b} K(x,t)\,dt &\leq 1, \quad a \leq x, y \leq b, \\
0 &\leq K(x,y) \leq k \quad a \leq x, y \leq b.
\end{align*}
\]

In the proof only values of the independent variables which lie in the interval \( (ab) \) are considered. We will make considerable use of the idea contained in the following two lemmas.

**Lemma 3.1:** If for \( x \) in \( S \) and \( y \) in \( S \)

\[
\int_{S} K(x,t)K(t,y)\,dt = K(x,y)
\]

and if \( S' = S - A \) with \( m(A) = 0 \), then

\[
\int_{S'} K(x,t)K(t,y)\,dt = K(x,y)
\]

for \( x \) in \( S' \) and \( y \) in \( S' \).

**Proof:** The proof is obvious since the first integral equals the second plus an integral over the set \( A \).
Since $A$ is of measure zero, the integral over $A$ is zero so the first and second integrals are equal.

**Lemma 3.2:** If over a set $A$, $f(t) > 0$ and $g(t) \geq 0$, $f(t)$ and $g(t)$ measurable, and if

$$\int_A f(t)g(t)\,dt = 0,$$

then

$$\int_A g(t)\,dt = 0.$$

**Proof:** The lemma is certainly true if $m(A) = 0$ so we assume that $m(A) > 0$. In any case

$$f(t)g(t) = 0$$

except on a set of measure zero, and since $f(t) > 0$, also $g(t) = 0$ except on this same set of zero measure. Thus

$$\int_A g(t)\,dt = 0.$$

Suppose now that for some $y_1$, $K(x,y_1) \equiv 0$ for almost all $x$, then

$$K(x,y_1) = \int_a^b K(x,t)K(t,y)\,dt \equiv 0$$

since the set on which the integrand is positive has zero measure. Let $F$ be the set of all $y$ such that

$$K(x,y) \equiv 0, \quad y \text{ in } F.$$
(This proves part (d) of Theorem 3.1. If the complement of \( F \) with respect to the interval \((ab)\) has measure zero, then the theorem is trivially true. We assume then that \( m[(ab) - F] > 0 \). Then by (25) if \( y \) is not in \( F \), \( K(x,y) > 0 \) on a set of positive measure so that

\[
\int_a^b K(x,y)dx > 0, \quad \text{y not in} \ F.
\]

**Lemma 3.3:** For almost all \( x \) not in \( F \)

\[
\int_a^b K(x,t)dt = 1.
\]

**Proof:** We define

\[
\alpha(x) = 1 - \int_a^b K(x,t)dt.
\]

Since \( K(x,t) \) is a kernel \( \alpha(x) \) is integrable. Moreover

\[
0 \leq \alpha(x) \leq 1.
\]

From (22) we have

\[
\int_a^b K(x,y)dy = \int_a^b \int_a^b K(x,t)K(t,y)dt dy = \int_a^b K(x,t)\int_a^b K(t,y)dy dt = \int_a^b K(x,t)[1-\alpha(t)]dt
\]

\[
= \int_a^b K(x,t)dt - \int_a^b \alpha(t)K(x,t)dt
\]

whence

\[
\int_a^b K(x,t)\alpha(t)dt \geq 0.
\]

Thus if \( B \) is the set of \( t \) on which \( \alpha(t) > 0 \), then
\( (27) \quad \int_{B \setminus F} K(x,t)\alpha(t)dt = 0 \)

where \( \setminus \) is the complement of \( F \) (with respect to the interval \((a, b)\)).

From (27) and Lemma 3.2 we can conclude that

\( (28) \quad \int_{B \setminus F} K(x,t)dt = 0 \)

since \( K(x,t) \geq 0 \) and \( \alpha(t) > 0 \) on \( B \). Integrating the identity (28) from \( a \) to \( b \) we get

\[ 0 = \int_{a}^{b} \left( \int_{B \setminus F} K(x,t)dt \right) dx = \int_{a}^{b} \left( \int_{B \setminus F} K(x,t)dx \right) dt. \]

On \( B \setminus F \) the integral \( \int_{a}^{b} K(x,t)dx > 0 \) by (26) so we must have \( m(B \setminus F) = 0 \). Thus the set of \( x \) not in \( F \) on which \( \alpha(x) > 0 \) has measure zero, which proves the lemma.

Now we define a new set \( S_1 \) by the relation

\[ S_1 = \text{interval } (ab) - (F + B \setminus F). \]

By (22), Lemmas 3.1 and 3.3 and the definition of the set \( F \) we have

\( (29) \quad \int_{S_1} K(x,t)K(t,y)dt = K(x,y), \quad x \text{ and } y \text{ in } S_1. \)

Moreover, since \( S_1 \) contains no \( x \) for which \( \alpha(x) > 0 \)

\( (30) \quad \int_{S_1} K(x,t)dt = 1, \quad x \text{ in } S_1 \)

and of course we still have
LEMMA 3.4: If

\( K(x_1, y) = K(x_2, y), \quad x_1 \text{ and } x_2 \text{ in } S_1 \)

for almost all \( y \text{ in } S_1 \), then

\( K(x_1, y) = K(x_2, y), \quad y \text{ in } S_1. \)

Proof: The proof is obvious since by (29)

\[
K(x_1, y) = \int_S K(x_1, t)K(t, y)dt
\]

and

\[
K(x_2, y) = \int_S K(x_2, t)K(t, y)dt,
\]

the two integrands differing on at most a set of measure zero.

If now for every \( x_1 \) and \( x_2 \) in \( S_1 \)

\( K(x_1, y) = K(x_2, y), \quad y \text{ in } S_1, \)

then part (a) of the theorem is true (there being in this event only one \( A_1 \)). Suppose then that there is an \( x_1 \) and an \( x_2 \) in \( S_1 \) such that

\( K(x_1, y) \neq K(x_2, y). \)

With this \( x_1 \) and \( x_2 \) we define the function

\( \beta(y) = K(x_1, y) - K(x_2, y), \)

and the three sets \( G^+, G^-, G^0 \) on which \( \beta(y) > 0, \beta(y) < 0 \) and \( \beta(y) = 0 \) respectively. (We are interested only in values of \( y \) in \( S_1 \) so the sets \( G^+, G^-, \) and \( G^0 \) will be taken only on \( S_1 \). Thus \( G^+ + G^- + G^0 = S_1 \).)
We can now prove

**Lemma 3.5**: \( m(G^+) > 0 \) and \( m(G^-) > 0 \).

**Proof**: Certainly \( m(G^+ + G^-) > 0 \) since by its definition \( \beta(y) \neq 0 \) and by Lemma 3.4 we see that \( \beta(y) \) must be different from zero on a set of positive measure. Thus at least either \( G^+ \) or \( G^- \) has positive measure. Suppose \( m(G^+) > 0 \) but \( m(G^-) = 0 \). Then \( \beta(y) \geq 0 \) for almost all \( y \) in \( S_1 \) and in fact

\[
\int_{S_1} \beta(y) dy = \int_{G^+} \beta(y) dy > 0.
\]

From the definition of \( \beta(y) \) and (30), however, we notice that

\[
\int_{S_1} \beta(y) dy = \int_{S_1} K(x_1, y) dy - \int_{S_1} K(x_2, y) dy = 0.
\]

From this contradiction we conclude that \( m(G^-) > 0 \) if \( m(G^+) > 0 \). In a similar manner we can show that \( m(G^+) > 0 \) if \( m(G^-) > 0 \) and since at least one of \( G^+ \) or \( G^- \) must have positive measure the lemma follows.

**Lemma 3.6**:

\[
\int_{cG^-} K(x, y) dy = 0, \text{ for almost all } x \text{ in } G^-.
\]

\[
\int_{cG^+} K(x, y) dy = 0, \text{ for almost all } x \text{ in } G^+.
\]

**Proof**: We first note that \( \beta(y) \) satisfies the relation

\[
\beta(y) = \int_{S_1} \beta(x) K(x, y) dx
\]
which is obvious from its definition and (29). From this we have

\[
\int_{G^+} \beta(y) dy = \int_{G^+} \int_{S_1} \beta(x) K(x,y) \, dx \, dy
\]

\[
= \int_{G^+} [\int_{G^+} \beta(x) K(x,y) \, dx + \int_{G^+} \beta(x) K(x,y) \, dx] \, dy
\]

\[
\leq \int_{G^+} \int_{G^+} \beta(x) K(x,y) \, dx \, dy = \int_{G^+} \beta(x) \int_{G^+} K(x,y) \, dy \, dx
\]

\[
\leq \int_{G^+} \beta(x) \left[ \int_{G^+} K(x,y) \, dy + \int_{G^+} K(x,y) \, dy \right] \, dx = \int_{G^+} \beta(x) \, dx,
\]

the last step by virtue of (30).

In the above chain the inequalities obviously can not hold so

\[
\int_{G^+} \beta(x) \left[ \int_{cG^+} K(x,y) \, dy \right] \, dx = 0,
\]

and since on $G^+$, $\beta(x) > 0$ we have by Lemma 3.2

(35) \[
\int_{G^+} \int_{cG^+} K(x,y) \, dy \, dx = 0.
\]

From this we see immediately that

\[
\int_{cG^+} K(x,y) \, dy = 0
\]

for almost all $x$ in $G^+$.

In a similar manner from (34) we can show that

(36) \[
\int_{G^-} \int_{cG^-} K(x,y) \, dy \, dx = 0
\]

from which the rest of the lemma follows.
By writing (35) and (36) in the forms

\[
\begin{align*}
\int_{G^+} \int_{G} K(x,y) \, dx \, dy &= 0 \\
\int_{G^-} \int_{G} K(x,y) \, dx \, dy &= 0
\end{align*}
\]

respectively, we can readily prove

**Lemma 3.7:**

\[
\begin{align*}
\int_{G^+} K(x,y) \, dx &= 0, \text{ for almost all } y \text{ in } G^+.
\int_{G^-} K(x,y) \, dx &= 0, \text{ for almost all } y \text{ in } G^-.
\end{align*}
\]

Now we disregard the exceptional sets of Lemmas 3.6 and 3.7 and denote by \(G^x, G^y,\) and \(G^z\) what is left of \(G^+, G^-,\) and \(G^0\) respectively. Also let

\[
S_1 = G^x + G^y + G^z.
\]

Since the exceptional sets were all of measure zero, \(m(S_1 - S_1) = 0\) so by (29) and Lemma 3.1

(37) \[
\int_{S_1} K(x,t)K(t,y) \, dt = K(x,y), \quad x \text{ and } y \text{ in } S_1.
\]

**Lemma 3.8:** If \(G'\) is any one of \(G^x, G^y,\) and \(G^z,\) then

\[
\int_{G'} K(x,t)K(t,y) \, dt = K(x,y), \quad x \text{ and } y \text{ in } G'.
\]

**Proof:** We write (37) in the form

(37') \[
K(x,y) = \int_{G^+} K(x,t)K(t,y) \, dt + \int_{G^-} K(x,t)K(t,y) \, dt + \int_{G^0} K(x,t)K(t,y) \, dt.
\]
If now \( x \) and \( y \) are both in \( \overline{G}^+ \), then the second and third integrals are zero by Lemma 3.6; similarly if \( x \) and \( y \) are both in \( \overline{G}^- \). On the other hand if \( x \) and \( y \) are both in \( \overline{G}^0 \), the first and second integrals are zero by Lemma 3.7. (Note that for the sets now under discussion there are no exceptions to the equations of Lemmas 3.6 and 3.7). This establishes Lemma 3.8.

Now over \( \overline{G}^+ \) or \( \overline{G}^- \) we have the same problem as before to consider, i.e.

\[
(38) \quad \int_{\overline{G}} K(x,t)K(t,y)\,dt = K(x,y), \quad x \text{ and } y \text{ in } \overline{G},
\]

\[
(39) \quad \int_{\overline{G}} K(x,t)\,dt = 1, \quad x \text{ in } \overline{G}
\]

where \( \overline{G} \) is either \( \overline{G}^+ \) or \( \overline{G}^- \). (Equation (39) is true by virtue of (30) and Lemma 3.6). Also of course

\[
0 \leq K(x,y) \leq k, \quad x \text{ and } y \text{ in } \overline{G}.
\]

Over \( \overline{G}^0 \) we have to consider the original problem; we still have (38) but in place of (39) we have

\[
(40) \quad \int_{\overline{G}^0} K(x,t)\,dt \leq 1, \quad \text{for } x \text{ in } \overline{G}^0.
\]

From (37) we see that if \( y \) is in \( \overline{G}^0 \), then

\[
K(x,y) = 0, \quad \text{for } x \text{ in } \overline{G}^+ \text{ or } \overline{G}^-,
\]

since by Lemma 3.7 the first and second integrals in (37') would be zero and the third would be zero by
Lemma 3.6. Now on $\mathfrak{F}^o$ there can be no set corresponding to $F$, since if for $y$ in $\mathfrak{F}^o$

$$K(x,y) = 0, \quad x \text{ in } \mathfrak{F}^o,$$

then by the above

$$K(x,y) = 0, \quad x \text{ in } \mathfrak{F}^o.$$

Thus

$$K(x,y) = \int_{\mathfrak{F}^o} K(x,t)K(t,y)dt + \int_{F} K(x,t)K(t,y)dt = 0 \quad \text{for all } x \text{ in the interval (ab). That is this } y \text{ is in } F \text{ which is impossible since } y \text{ is in } \mathfrak{F}^o \text{ which is contained in the complement of } F.$$

Merely by dropping from $\mathfrak{F}^o$ then, at most another set of measure zero (the one corresponding to $B_0 F$) we have (39) also holding on $\mathfrak{F}^o$ (or if $m(\mathfrak{F}^o) = 0$ we, of course, need no longer consider this set). If $m(\mathfrak{F}^o)>0$ we denote by $\mathfrak{F}''$ what is left after this set of measure zero is removed.

LEMMMA 3.9: If $x$ is in $C$ where $C$ is any one of $\mathfrak{F}^+, \mathfrak{F}^-$ and $\mathfrak{F}''$, then

$$\int_D K(x,y)dy = 0$$

where $D$ is any other one of the three sets different from the one containing $x$.

Proof: If $C$ is $\mathfrak{F}^+$ or $\mathfrak{F}^-$ this is obvious from Lemma 3.6.
If \( C = \overline{G} \), then since

\[
1 = \int_{\overline{G}} K(x,y) \, dy = \int_{G^+} K(x,y) \, dy + \int_{G^-} K(x,y) \, dy + \int_{\overline{G}} K(x,y) \, dy
\]

by (39), we must have

\[
\int_{G^+} K(x,y) \, dy + \int_{G^-} K(x,y) \, dy = 0.
\]

The lemma follows since \( K(x,y) \geq 0 \).

We have seen that by dropping a finite number of sets of measure zero, together comprising only a set of measure zero, we have broken the original set \( S_1 \) up into two or three sets \( A_i \) (\( i = 1, 2, 3 \)) over each of which we must consider the same problem as defined by (29), (30) and (31). Moreover, by Lemma 3.9

\[
\int_{A_i} K(x,y) \, dy = 0, \quad \text{for } x \text{ not in } A_i.
\]

If now for every \( x_1 \) and \( x_2 \) both in \( A_i \) (\( i = 1, 2, 3 \)) we have

\[
K(x_1,y) = K(x_2,y) \quad \text{for } y \text{ in } A_i,
\]

then the theorem is true. Suppose then that there is one of the \( A_i \) containing an \( x_1 \) and \( x_2 \) such that

\[
K(x_1,y) \nmid K(x_2,y), \quad y \text{ in } A_i.
\]

A repetition of the previous argument, applied instead
over $iA_1$ in place of $S_1$, shows that $iA_1$ can be broken up into three sets (possibly only two) like the original three $iA_1, iA_2, iA_3$, again dropping at most a set of measure zero. We denote the three sets so obtained by $sA_1, i = 1, 2, 3$. We would have

$$iA_1 = sA_1 + sA_2 + sA_3 + H$$

where $m(H) = 0$.

The above process can be repeated as long as there is one of the $A_i$'s containing an $x_1$ and $x_2$ such that (41) is true over this set. At the $n^{th}$ stage of the process we have

**LEMMA 3.10:** $m(sA_r) \leq (b - a) - \frac{2n}{k}$.

**Proof:** Since $|K(x,y)| \leq k$ we have

$$1 = \int_{A_r} K(x,t) dt \leq k \cdot m(sA_r), \quad r = 1, 2, 3,$$

whence

$$m(sA_r) \geq \frac{1}{k}.$$

Moreover, since

$$m(s_{-1}A_6) = m(sA_1) + m(sA_2) + m(sA_3) \geq m(sA_r) + \frac{2}{k},$$

we see that

$$m(sA_r) \leq m(s_{-1}A_1) - \frac{2}{k}.$$
From (42) we also get

\[
m(nA_t) \leq m(n-A_t) - \frac{2}{k} - \frac{2}{k}
\]

\[
= m(n-A_t) - 2(\frac{2}{k})
\]

and by induction

\[
m(nA_r) \leq m(S_1) - \frac{2n}{k}
\]

\[
\leq (b - a) - \frac{2}{k} n
\]

thus establishing the lemma.

From this lemma we immediately see that there is an upper bound to \( n \) since the measure of any of the \( A \)'s is positive. That is, eventually we must reach a stage at which for every \( x_1 \) and \( x_2 \) in the same \( A \) we must have

\[
K(x_1,y) = K(x_2,y), \quad y \text{ in } A.
\]

In fact, then

\[
K(x_1,y) = K(x_2,y)
\]

for all \( y \) since

\[
K(x_1,y) = \int_{A} K(x_1,t)K(t,y)dt + \int_{cA} K(x_1,t)K(t,y)dt
\]

and

\[
K(x_2,y) = \int_{A} K(x_2,t)K(t,y)dt + \int_{cA} K(x_2,t)K(t,y)dt.
\]

The integrals over \( cA \) are zero (\( cA \) includes \( F \), the
discarded sets of measure zero and for the rest we can appeal to Lemma 3.9), and the first integrals are the same in both cases. Thus

\[ K(x, y) = p_n(y) \]

for all \( x \)'s in the same \( A_n \) (which proves part (a) of the theorem).

By an argument similar to that on page 31 showing that there was no set in \( F^c \) corresponding to \( F \) we can prove part (b) of the theorem, that is

\[ p_n(y) > 0, \quad y \text{ in } A_n. \]

At each stage of the process of breaking the interval up into the sets \( F, A_1, A_2, \ldots \) we discarded a set of measure zero. This was done only a finite number of times so if all these sets are put into one set \( V \) we will have \( m(V) = 0 \), which proves part (e) of the theorem.

In this proof of the theorem we see that in the particular case of a bounded kernel with \( S \) a finite interval there can be only a finite number of the \( A_i \). Of course, this follows from the general theorem too.
IV. GENERAL LINEAR RECURRENCE RELATIONS BETWEEN
THE ITERATED KERNELS

We recall from Chapter I that

\[ B_n(x,y) = A_n K(x,y) - n \int_a^b K(x,t) B_{n-1}(t,y) \, dt \]  

and that

\[ A_{n+1} = \int_a^b B_n(t,t) \, dt. \]  

Suppose now that for some \( n \)

\[ B_n(x,y) \neq 0. \]

Then from (7) \( A_{n+1} = 0 \) and we have from (8)

\[ B_{n+1}(x,y) = A_{n+1} K(x,y) - (n+1) \int_a^b K(x,t) B_n(t,y) \, dt \]

\[ = 0 \]

etc. Thus the series for \( D(x,y; \lambda) \) terminates after \( n \)
terms if we have \( B_n(x,y) \neq 0 \). Moreover the series for
\( D(\lambda) \) terminates and both \( D(\lambda) \) and \( D(x,y; \lambda) \) are then polynomials in \( \lambda \).

It is known that if the kernel \( K(x,y) \) is separable,
then \( D(\lambda) \) is a polynomial in \( \lambda \). The converse is
not necessarily true, but with a few additional con-
ditions on \( D(x,y; \lambda) \) the necessity follows. We prove
first
THEOREM 4.1: If $K(x,y)$ is a bounded kernel and if
for some $p$

$$ a_1 K_1(x,y) + a_2 K_2(x,y) + \cdots + a_p K_p(x,y) = 0, \quad a_1 \neq 0, $$

for $x$ and $y$ in the square $a \leq x, y \leq b$,

then $K(x,y)$ is separable.

Proof: If (44) is satisfied with $a_2 = a_3 = \cdots = a = 0$, then $K(x,y) = K(x,y) = 0$ which is obviously separable. Suppose then that not all $a_i = 0$,

$i = 2, 3, \ldots, p$. Equation (44) can be written

$$ K(x,y) = \int_a^b H(x,t)K(t,y)dt $$

where

$$ H(x,t) = -\frac{1}{a_1} (a_2 K_1(x,t) + \cdots + a_p K_{p-1}(x,t)). $$

If $H(x,t) = 0$, then again, from (45) $K(x,y) = 0$, so we assume $H(x,t) \neq 0$. Since $K(x,y)$ is bounded, then $H(x,t)$ is likewise, for

$$ |H(x,t)| \leq \frac{1}{|a_1|} \left( |a_2| k + \cdots + |a_p| k^{p-1} (b-a)^{p-2} \right) \leq M, $$

where, as before, $|K(x,t)| \leq k$.

The proof now is similar to that of Theorem 2.2.

For any fixed $y_1$ in the interval $(ab)$ $K(x,y_1)$ is a solution of

$$ u(x) = \int_a^b H(x,t)u(t)dt, $$
and from the linearly independent non-trivial $K(x,y_1)$ a set of orthonormal solutions $u_i(x)$, $i = 1, 2, \ldots, n$, can be constructed. Again we find that $n$ is bounded since by Bessel's inequality

$$\sum_{i=1}^{n} [u_i(x)]^2 = \sum_{i=1}^{n} \left[ \int_{a}^{b} H(x,t)u_i(t)dt \right]^2 \leq \int_{a}^{b} \left[ H(x,t) \right]^2 dt$$

whence

$$n = \sum_{i=1}^{n} \int_{a}^{b} [u_i(x)]^2 dx \leq \int_{a}^{b} \left[ H(x,t) \right]^2 dt dx \leq M^2(b-a)^2.$$  

If $N$ is the number of linearly independent $K(x,y_1)$ and $u_i(x)$, $i = 1, 2, \ldots, N$, the corresponding orthonormal system of functions, we have for any other $y$

$$(46) \quad K(x,y) = \sum_{i=1}^{N} u_i(x)v_i(y).$$

Thus the theorem is proved.

From this expansion we can deduce some other interesting results. Let us define the two vectors

$\mathbf{U}(x) = (u_1(x), u_2(x), \ldots, u_N(x))$

$\mathbf{V}(y) = (v_1(y), v_2(y), \ldots, v_N(y))$.

We can then write

$$K(x,y) = \mathbf{U}(x)\mathbf{V}^T(y)$$

where $\mathbf{V}^T(y)$ is the transpose of $\mathbf{V}(y)$. Furthermore
\[ K_2(x,y) = \int_a^b K(x,t)K(t,y)\,dt = \int_a^b U(x)V^T(t)U(t)V^T(y)\,dt \]
\[ = U(x) \left( \int_a^b V^T(t)U(t)\,dt \right) V^T(y) \]
\[ = U(x)CV^T(y) \]

where \( C \) is the square matrix whose elements are

\[ c_{ij} = \int_a^b v_i(t)u_j(t)\,dt, \quad i,j = 1,2,\ldots,N. \]

Similarly,

\[ K_3(x,y) = U(x)C^2V^T(y) \]

and in general

\[ K_n(x,y) = U(x)C^{n-1}V^T(y). \]

Since it is supposed that the iterated kernels satisfy the relation (44) we have

\[ \sum_{i=1}^p a_i K_1(x,y) = \sum_{i=1}^p a_i U(x)C^{i-1}V^T(y) \]
\[ = U(x) \left( \sum_{i=1}^p a_i C^{i-1} \right) V^T(y) = 0. \]

The elements of \( U(x) \) are linearly independent and if the elements of \( V(y) \) are also linearly independent, then (48) is equivalent to the matrix equation

\[ \sum_{i=0}^p a_i C^{i-1} = 0. \]

That the elements of \( V(y) \) are linearly independent
can be shown as follows. We state this as

**Lemma 4.1:** In the expansion (22) the \( v_1(y) \) are

**Linearly independent.**

**Proof:** We have

\[
K(x, y) = \sum_{i=1}^{N} v_i(y) u_i(x)
\]

whence, as in the proof of Theorem 2.3,

\[
\int_{a}^{b} K(x, y) u_i(x) \, dx = \sum_{j=1}^{N} v_j(y) \delta_{ij} = v_i(y).
\]

Suppose now that there are constants \( b_1, b_2, \ldots, b_N \)

such that

\[
\sum_{i=1}^{N} b_i v_i(y) = 0.
\]

We would have then

\[
\sum_{i=1}^{N} b_i \int_{a}^{b} u_i(x) K(x, y) \, dx = \int_{a}^{b} f(x) K(x, y) \, dx = 0
\]

where

\[
f(x) = \sum_{i=1}^{N} b_i u_i(x).
\]

Also for all integral \( n \geq 2 \) we have, therefore,

\[
\int_{a}^{b} f(x) K(x, y) \, dx \, K_{n-1}(y, t) \, dy = \int_{a}^{b} f(x) \left( \int_{a}^{b} K(x, y) K_{n-1}(y, t) \, dy \right) \, dx
\]

\[
= \int_{a}^{b} f(x) K_n(x, t) \, dx = 0.
\]

Thus

\[
\sum_{n=2}^{p} \frac{a_n}{a_1} \int_{a}^{b} f(x) K_{n-1}(x, t) \, dx = \int_{a}^{b} f(x) H(x, t) \, dx = 0.
\]

Then for \( i = 1, 2, \ldots, N \),
\[
0 = \int_{a}^{b} u_1(t) \int_{a}^{b} f(x) H(x,t) dx dt = \int_{a}^{b} f(x) \int_{a}^{b} H(x,t) u_1(t) dt dx
= \int_{a}^{b} \sum_{j=1}^{N} b_j u_j(x) u_1(x) dx = \sum_{j=1}^{N} b_j \delta_{ij} = b_i.
\]

That is
\[
\sum_{i=1}^{N} b_i v_1(y) = 0
\]

if and only if all the \( b_i = 0 \). Thus the \( v_1(y) \) are linearly independent.

With this lemma we have then proven

**Theorem 4.2**: If a bounded kernel satisfies the relation
\[
a_1 K_1(x,y) + \ldots + a_p K_p(x,y) = 0
\]
with \( a_1 \neq 0 \) for \( a < x, y < b \), then the matrix \( C \) as defined in (47) satisfies the equation
\[
a_1 I + a_2 C + \ldots + a_p C^{p-1} = 0.
\]

**Corollary 4.2.1**: A bounded idempotent kernel can be written as
\[
K(x,y) = \sum_{i=1}^{N} u_1(x)v_1(y)
\]

with
\[
\int_{a}^{b} u_1(t) u_j(t) dt = \int_{a}^{b} u_1(t) v_j(t) dt = \delta_{ij}.
\]

**Proof**: Such a kernel satisfies (44) with \( p = 2 \) and
\( a_2 = -a_1 = 1 \). By Theorem 4.2 the matrix then satisfies
\[-I + C = 0\]
so that
\[ C = I. \]
That is
\[ c_{ij} = \int_a^b v_i(t)u_j(t)dt = \delta_{ij}. \]
(Of course, it is presumed here that the \( u_i(x) \) are already chosen orthonormal as in Theorem 4.1 so the above completes the proof).

A partial converse of Theorem 4.1 is easily shown. Thus let \( K(x,y) \) be a bounded separable kernel
\[ K(x,y) = \sum_{i=1}^{N} u_i(x)v_i(y) = U(x)V^T(y) \]
and \( C \) the matrix defined in (47). (Here the \( u_i(x) \) and \( v_i(y) \) may or may not be linearly independent).

We can easily prove

**Theorem 4.3:** For a bounded separable kernel there is a linear combination of the iterated kernels which is identically zero.

**Proof:** Indeed
\[ \sum_{i=1}^{N+1} a_i K_i(x,y) = U(x) \left( \sum_{i=1}^{N+1} a_i c^{i-1} \right) V^T(y) \equiv 0 \]
if the polynomial
since every matrix "satisfies" its own characteristic equation.

We note here that the linear combination might be of smaller "order" than N+1 if the minimum equation of the matrix C is of lesser degree than N or if the \( u_1(x) \) or \( v_1(y) \) are linearly dependent.

The above results enable us to prove

**THEOREM 4.4:** If for a bounded kernel \( K(x,y) \), \( D(x,y; \lambda) \) is a polynomial in \( \lambda \) and \( D(\lambda) \) is a polynomial of the same degree, then \( K(x,y) \) is separable.

**Proof:** Let \( D(x,y; \lambda) \) be a polynomial in \( \lambda \) of degree \( n \). Then

\[
D(x,y; \lambda) = \sum_{i=1}^{N+1} a_i \lambda^{i-1} = \det[C - \lambda I]
\]

Since \( D(\lambda) \) is of the same degree as \( D(x,y; \lambda) \) we have

\[
B_n(x,y) = 0, \quad \text{but} \quad B_{n-1}(x,y) \neq 0.
\]

From (50) and Theorem 1.3 we see that

\[
A_{n+1} = 0, \quad \text{but} \quad A_n = \int_a^b B_{n-1}(x,x)dx \neq 0.
\]

\[
A_n K_1(x,y) = n A_{n-1} K_2(x,y) + n(n-1) A_{n-2} K_3(x,y) + \cdots + (-1)^n n! K_{n-1}(x,y) = 0
\]

with \( A_n \neq 0 \) by (51). Thus the iterated kernels satisfy a linear relation as in Theorem 4.1. The other
conditions of this theorem are also fulfilled so we conclude that $K(x,y)$ is separable.

In general it is not sufficient for $D(x,y; \lambda)$ and $D(\lambda)$ merely to be polynomials in order for $K(x,y)$ to be separable, as we will show by some examples.

We first prove

**Lemma 4.2:** If $\Psi_i(x)$, $i = 1,2,\ldots$ are an orthogonal system of functions on the interval $(ab)$, if $\phi_i(y) \neq 0$, $i = 1,2,\ldots$ are a set of linearly independent functions and if

$$K(x,y) = \sum_{i=1}^{\infty} \Psi_i(x)\phi_i(y)$$

converges uniformly in $x$ in the interval $(ab)$, then $K(x,y)$ is not separable.

**Proof:** Suppose we have

$$\sum_{i=1}^{\infty} \Psi_i(x)\phi_i(y) = K(x,y) = \sum_{i=1}^{N} u_i(x)v_i(y).$$

Since the infinite series converges uniformly we can multiply by $\Psi_j(x)$ and integrate from $a$ to $b$ so that

$$\int_a^b [\Psi_j(x)]^2 dx \phi_j(y) = \sum_{i=1}^{N} v_i(y) \int_a^b u_i(x)\Psi_j(x) dx$$

or

$$\phi_j(y) = \sum_{i=1}^{N} a_{ij}v_i(y) \quad j = 1,2,\ldots.$$

This, however, is impossible since if each $\phi_j(y)$ were
expressible as a linear combination of the $N$ functions $v_1(y)$, there could be no more than $N$ linearly independent $\phi_j(y)$. From this contradiction we conclude that $K(x,y)$ can not be written as a finite sum
\[ \sum_{i=1}^{N} u_i(x)v_i(y), \text{ i.e. } K(x,y) \text{ is not separable.} \]
Consider now the kernel defined by
\[ K(x,y) = \sum_{n=1}^{\infty} \frac{1}{n^2} \sin(2n-1)\pi x \cdot \sin 2ny \]
and take the interval $(ab)$ to be $(01)$. The series defining this $K(x,y)$ is obviously uniformly convergent so we have
\[ K_{1}(x,y) = \sum_{n,m=1}^{\infty} \frac{1}{n^2m^2} \sin(2n-1)\pi x \cdot \sin 2ny \int_{0}^{1} \sin((2n-1)\pi t) \cdot \sin 2\pi nt \cdot dt \]
\[ = 0, \]
and consequently
\[ K_i(x,y) = 0, \quad i \geq 2. \]
Also
\[ A_1 = \int_{0}^{1} K(t,t)dt = \sum_{n=1}^{\infty} \frac{1}{n^2} \int_{0}^{1} \sin((2n-1)\pi t) \cdot \sin 2\pi nt \cdot dt \]
\[ = 0. \]
Thus from (9)
\[ B_1(x,y) = A_1K_1(x,y) - A_0K_2(x,y) = 0, \]
and consequently

$$A_n = \int_0^1 B_1(x,x)dx = 0,$$

e tc. We have then

$$D(x,y; \lambda) = \lambda K(x,y),$$
a polynomial in \(\lambda\) of the first degree, and

$$D(\lambda) = 1$$
a polynomial of the zeroth degree.

This kernel is bounded, and although by Lemma 4.2 it is not separable, the iterated kernels do satisfy a linear relation

$$a_1K_1(x,y) + a_2K_2(x,y) + \cdots + a_pK_p(x,y) = 0,$$

however, only if \(a_1 = 0\).

Another example is the continuous kernel

$$K(x,y) = y-x, \quad 1 \geq y \geq x \geq 0$$

$$= 0, \quad 0 \leq y \leq x \leq 1.$$ 

Calculating the iterated kernels we have first

$$K_1(x,y) = \int_0^1 K(x,t)K(t,y)dt$$

$$= 0, \quad y < x$$

$$= \int_x^y (t-x)(y-t)dt = \frac{(y-x)^3}{3}, \quad y \geq x,$$

and in general
\[ K_n(x, y) = 0, \quad y < x \]
\[ \frac{(y-x)^{2n-1}}{(2n-1)!}, \quad y \geq x. \]

Thus \( p_n = \int_0^1 K_n(x, x) dx = 0 \) whence it follows from equation (10) that \( \lambda_n = 0, \ n \geq 1 \). However, we have from (9) that \( B_n(x, y) = (-1)^n n! K_{n+1}(x, y) \neq 0 \), so for this kernel
\[ D(\lambda) = 1 \]

while \( D(x, y; \lambda) = \lambda K(x, y) + \lambda^2 K_2(x, y) + \cdots + \lambda^n K_n(x, y) + \cdots \)

an infinite series in \( \lambda \).

We have seen that there are non-separable kernels having one or both of \( D(x, y; \lambda) \) and \( D(\lambda) \) as polynomials, but if in addition \( K(x, y) \) is continuous and symmetric, i.e. if
\[ K(x, y) = K(y, x), \]
then \( D(x, y; \lambda) \) and \( D(\lambda) \) are polynomials only if \( K(x, y) \) is separable. We prove this in

THEOREM 4.5: If \( K(x, y) \) is a bounded continuous symmetric kernel and if \( D(\lambda) \) is a polynomial, then \( K(x, y) \) is separable.

Proof: We have from Corollary 1.3.1 (equation (10))
(52) \[ A_{n+r} = \int_a^b \left[ A_{n+r-1}K(x,x) - (n+r-1)A_{n+r-2}K_1(x,x) + \cdots \right. \\
\left. + (-1)^{n+r-1}(n+r-1)!K_{n+r}(x,x) \right] \, dx \]

Suppose \( D(\lambda) \) is a polynomial of degree \( n \). Then

\[ A_n \neq 0, \quad \text{but} \quad A_{n+r} = 0, \quad r = 1, 2, \ldots \]

Thus from equation (52)

\[ 0 = \int_a^b \left[ (-1)^{r+1}(n+r-1)\cdots(n+1)A_nK_r(x,x) + \cdots \right. \\
\left. + (-1)^{n+r-1}(n+r-1)!K_{n+r}(x,x) \right] \, dx \]

\[ = (-1)^{r+1}(n+r-1)\cdots(n+1)\int_a^b \int_a^b \left[ a_1K_1(x,t) + \cdots \right. \\
\left. + a_{n+1}K_{n+1}(x,t) \right] K_{r-1}(t,x) \, dt \, dx \]

\[ r = 2, 3, \ldots \]

where

\[ a_p = \frac{n!}{(n-p+1)!} \frac{A_{n-p+1}}{A_n} \quad p = 1, 2, \ldots, n+1. \]

Furthermore

\[ a_1 = A_n \neq 0. \]

Since now \( K(x,t) \) is symmetric we can write this as

\[ \int_a^b \int_a^b \left[ a_1K_1(x,t) + \cdots + a_{n+1}K_{n+1}(x,t) \right] K_r(x,t) \, dt \, dx = 0 \]

\[ r = 1, 2, \ldots \]
Therefore,
\[
\sum_{r=1}^{n+1} a_r \int_a^b \left[ a_1 K_1(x,t) + \cdots + a_{n+1} K_{n+1}(x,t) \right] K_r(x,t) dx = 0,
\]
whence, since \( K(x,t) \) is continuous,
\[
a_1 K_1(x,t) + \cdots + a_{n+1} K_{n+1}(x,t) = 0
\]
for \( a \leq x, y \leq b \). By Theorem 4.2 then, \( K(x,t) \) is separable.

**Corollary 4.5.1:** If \( K(x,y) \) is a bounded continuous symmetric kernel and if \( D(x,y; \lambda) \) is a polynomial in \( \lambda \), then \( K(x,y) \) is separable.

**Proof:** Since \( D(x,y; \lambda) \) is a polynomial we have for some \( n \)
\[
B_{n+r}(x,y) = 0, \quad r = 0, 1, 2, \ldots
\]
Then
\[
A_{n+r+1} = \int_a^b B_{n+r}(x,x) dx = 0, \quad r = 0, 1, 2, \ldots
\]
Thus if \( D(x,y; \lambda) \) is of degree \( n \) in \( \lambda \), \( D(\lambda) \) is a polynomial of at most degree \( n \) and by Theorem 4.5 \( K(x,y) \) is separable.
V. SUMMARY

The principle result obtained in this thesis is the theorem that if the iterated kernels of an integral equation satisfy a linear relation

\[ a_1 K_1(x, y) + a_2 K_2(x, y) + \cdots + a_n K_n(x, y) = 0, \quad a_1 \neq 0, \]

then the kernel \( K(x, y) \) must be of the special form

\[ \sum_{i=1}^{N} u_i(x)v_i(y). \]

Using this result it is shown that only such kernels can have a Fredholm determinant \( D(\lambda) \) and Fredholm first minor \( D(x, y; \lambda) \) being polynomials in \( \lambda \) of the same degree. Also in the particular case of a continuous symmetric kernel it is shown that if either \( D(\lambda) \) or \( D(x, y; \lambda) \) are polynomials in \( \lambda \) that then \( K(x, y) \) must be of the special form given above.

Further properties of idempotent kernels, i.e. ones for which

\[ K_n(x, y) = K_1(x, y), \]

are deduced. The connection between such kernels and an idempotent Markoff process is pointed out thus indicating a possible application of the theory of this thesis to certain problems in probability. An alternate proof of a known result concerning such Markoff processes is given.

The results of the thesis are generally derived under the assumption of measurability and boundedness of the kernel \( K(x, y) \). Also integration is in the sense of Lebesgue.
VI. ACKNOWLEDGMENT

The author would like gratefully to acknowledge the encouragement and helpful suggestions so generously offered by Dr. Henry P. Thielman during the presentation of this thesis.

The help of Dr. George W. Brown and his suggestion of an area of work which ultimately led to this thesis is also greatly appreciated.