

## A PROBLEM IN CONVEXITY LEADING TO THE ANALYSIS OF TWO FUNCTIONAL EQUATIONS

JOHN V. RYFF

**ABSTRACT.** Transformation semigroups can often be studied effectively by examining their orbit structure. If the class of transformations has a special quality, such as convexity, it is generally reflected in the orbits. This work is concerned with such a circumstance. The goal is to examine the behavior of transformations on extreme points of orbits through the construction of a class of extreme operators. The construction leads naturally to the study of two functional equations which are analyzed in detail. Information about solutions is obtained through different  $L^2$ -methods depending on whether or not two basic parameters are rational or irrational. In two cases all solutions are classified. In a third an example of a spanning set of solutions is obtained. Techniques of harmonic analysis and ergodic theory are used to study the functional equations.

Functional equations often as not are studied because their solutions represent intermediate steps in a chain of events leading to the resolution of some other problem. One would like to accomplish this step efficiently, and efforts are undertaken to put the equation in some context in which it may be solved by existing methods. On occasion this approach fails and an investigator is left to his own resources to find a solution. A case in point is the equation

$$(e) \quad af(ax) + bf(bx + a) = bf(bx) + af(ax + b)$$

and its companion equation

$$(E) \quad F(ax) + F(bx + a) = F(bx) + F(ax + b).$$

For purposes of solving the original problem, we make the further assumptions that  $0 < a < b < 1$  and  $a + b = 1$ .

The genesis of (e) is a problem concerning convex sets of linear operators. If  $K$  is such a set and

$$\Omega(x) = \{Tx : T \in K\}$$

is called the *orbit* of  $x$ , then we can ask whether extreme points of  $K$  and extreme points of  $\Omega(x)$  are in any way related. For example, let  $K$  be the class of linear operators  $T: L^1(0, 1) \rightarrow L^1(0, 1)$  which satisfy

- (i)  $T \geq 0$ ,
- (ii)  $T1 = 1$ ,
- (iii)  $\int Tf = \int f$ .

---

Received by the editors November 25, 1986.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 46A55; Secondary 39B05.

*Key words and phrases.* Functional equations, convexity, extreme points, extreme operators, doubly stochastic, ergodic mappings, measure-preserving transformations.

It is understood that (i) means  $Tf \geq 0$  if  $f \geq 0$ , while in (ii), 1 denotes the constant function. This class has been studied extensively [5, 8, 9]. It is clearly convex and is also compact in a sufficiently weak topology [4]. As operators, the extreme points of  $K$  are difficult, perhaps impossible to classify (but see [2] for a different characterization). Nevertheless, for each element  $f$  of  $L^1(0, 1)$ , the extreme points of  $\Omega(f)$  are known. One can then ask whether an extreme element of  $K$  must preserve extreme points of some  $\Omega(f)$ . The question must be qualified slightly since  $T1 = 1$  implies that  $\Omega(1)$  is trivial.

The extreme points of  $\Omega(f)$  are simply the functions  $g$  which are equimeasurable with  $f$  (i.e., have the same distribution). If we denote this by writing  $g \sim f$ , then our question is: If  $T$  is extreme, does there exist a nonconstant  $f$  in  $L^1(0, 1)$  and a  $g \sim f$  such that  $Tf = g$ ? That is, for some orbit does  $T$  carry one extreme point onto another? As it stands, the problem is not easily managed. But a result of Sakai and Shimogaki [8] simplifies the search considerably, for if  $Tf = g$ , where  $f$  and  $g$  are equimeasurable, then  $T\chi_{E_\lambda} = \chi_{F_\lambda}$  for each of the level sets  $E_\lambda = \{f > \lambda\}$ ,  $F_\lambda = \{g > \lambda\}$ . The measures of these sets must be the same. Therefore  $\chi_{F_\lambda}$  is an extreme point of  $\Omega(\chi_{E_\lambda})$ —and vice versa. Our problem is then reduced to deciding whether an extreme operator can be constructed which fails to preserve characteristic functions other than the constant functions 0 and 1. An example producing such unpleasant extreme operators can be constructed in the following way. Choose two measure-preserving transformations  $\phi$  and  $\psi$  of  $[0, 1]$ . The composition operators  $T_\phi$  and  $T_\psi$  defined by  $T_\phi f = f \circ \phi$ ,  $T_\psi f = f \circ \psi$  are extreme points of  $K$  [4]. Construct a new element of  $K$  by selecting  $g \in L^\infty$ ,  $0 \leq g \leq 1$ , and setting

$$(1) \quad (Tf)(x) = g(x)(T_\phi f)(x) + (1 - g(x))(T_\psi f)(x).$$

This sliding convex combination must be so arranged that  $T$  is an extreme point of  $K$  which does not preserve characteristic functions. The requirement that  $T \in K$  will be satisfied once condition (iii) has been established:

$$\int Tf = \int (gT_\phi f + (1 - g)T_\psi f) = \int f.$$

Taking adjoints and noting that  $T^*1 = 1$ , this becomes

$$\int (T_\phi^*g - T_\psi^*g)f = 0.$$

As this must be true for all  $f \in L^1$ ,  $g$  must be chosen so that

$$(2) \quad T_\phi^*g = T_\psi^*g.$$

In general, the adjoint of a composition operator cannot be easily displayed. If, however,  $\phi$  is an invertible measure-preserving transformation, then  $T_\phi^*$  is just  $T_{\phi^{-1}}$ . A consequence of this is the observation that if one of the two measure-preserving transformations is invertible, any  $T \in K$  defined by (1) must carry a characteristic function into a characteristic function. If  $\phi$  is invertible then, by (2), any admissible  $g$  would have to satisfy

$$g \circ \phi^{-1} = T_\psi^*g.$$

For each  $G_\lambda = \{g > \lambda\}$  we must have

$$T_\psi^*\chi_{G_\lambda} = \chi_{\phi(G_\lambda)}$$

so that  $T_\psi T_\psi^* \chi_{G_\lambda} = \chi_{\psi^{-1}\phi(G_\lambda)}$ . But  $T_\psi T_\psi^* \chi_{G_\lambda} = \chi_{G_\lambda}$ . If  $F_\lambda = \phi(G_\lambda)$  we arrive at  $\phi^{-1}(F_\lambda) = \psi^{-1}(F_\lambda)$  or  $T\chi_{F_\lambda} = \chi_{F_\lambda}$ . In the event that  $G_\lambda$  is either  $[0, 1]$  or  $\emptyset$  for all  $\lambda$ , this would imply that  $g$  is a constant and therefore that  $T$  is either not extreme or else  $T$  reduces to  $T_\phi$  or  $T_\psi$ .

To find an interesting case it is necessary to assume that  $\phi$  and  $\psi$  are not invertible. Among the simplest examples which come to mind are those mappings which are similar to the standard example  $x \rightarrow 2x \pmod{1}$ . Let  $a$  and  $b$  be any numbers in the open interval  $(0, 1)$  which satisfy  $a + b = 1$ . Define two functions

$$\phi(x) = \begin{cases} a^{-1}x, & 0 \leq x \leq a, \\ b^{-1}(x - a), & a < x \leq 1, \end{cases} \quad \psi(x) = \begin{cases} b^{-1}x, & 0 \leq x \leq b, \\ a^{-1}(x - b), & b < x \leq 1. \end{cases}$$

These are measure preserving and lead to nontrivial operators (1) provided that  $g$  satisfies (2). In this particular instance (2) becomes

$$ag(ax) + bg(bx + a) = bg(bx) + ag(ax + b).$$

A number of results concerning solutions to this equation may be found in [1, 6, and 7]. The purpose of this work is to give a relatively complete analysis of solutions to (e) and, to a lesser extent, (E) in an  $L^2$ -setting.

We separate the cases where  $a$  and (hence)  $b$  are rational or irrational. It is important to note at the outset that operators of the form (1) will be extreme in the class  $K$  if and only if the corresponding solution  $g$  of (e) is extreme among all solutions which satisfy  $0 \leq g \leq 1$  [7, §4].

**1. Extension.** An extension result will first be given which plays a fundamental role in what is to follow. It also shows why harmonic analysis can be expected to enter into the discussion. For if  $0 < x < 1$ , then  $ax < bx < bx + a < ax + b$  always holds so that any solution  $f$  of (e) satisfies

$$(3) \quad f(ax + b) = \frac{b}{a}(f(bx + a) - f(bx)) + f(ax).$$

In other words,  $f$  is prescribed once it is known for the three lesser values. In particular, if  $f$  is given on  $[0, b)$ , a solution is generated by repeated applications of (3). A similar argument applies to equation (E) with the term  $b/a$  replaced by unity. This is all worked out in some detail in [7].

What is important is the observation that this technique can be used to extend solutions of (e) and (E) beyond  $[0, 1]$ . In general, they will not be unique. However, there is a natural choice which agrees with obvious extensions for smooth (e.g. analytic) solutions defined on  $[0, 1]$ . For (E) one takes

$$F(bx + a) = F(bx) + F(b + a(x - 1)) - F(a(x - 1)) - (F(a) + F(b) + F(0) + F(1)),$$

and

$$bf(bx + a) = bf(bx) + af(b + a(x - 1)) - af(a(x - 1))$$

for (e). The first equation is only valid for solutions of (E) which are continuous on  $[0, 1]$ . Both hold only for  $1 \leq bx + a \leq 1 + a$ , but once an extension has been obtained this far, a unique continuation takes place via the functional equation (except that now  $ax + b < bx + a$ ). Moreover, we can eliminate the constant term  $F(a) + F(b) + F(0) + F(1)$  in the first equation above by adding to  $F$  a

suitable multiple of the particular solution  $\Phi_0$  given by  $\Phi_0(x) = x^2 - x$ , since  $\Phi_0(a) + \Phi_0(b) + \Phi_0(0) + \Phi_0(1) = -2ab \neq 0$ . With that in place, we rewrite the extension equations as

$$\begin{aligned} F(bx + a) - F(bx) &= F(a(x - 1) + b) - F(a(x - 1)) \\ &= F(b(x - 1) + a) - F(b(x - 1)) \end{aligned}$$

and

$$\begin{aligned} b(f(bx + a) - f(bx)) &= a(f(a(x - 1) + b) - f(a(x - 1))) \\ &= b(f(b(x - 1) + a) - f(b(x - 1))) \end{aligned}$$

because in the range  $1 \leq bx + a \leq 1 + a$  one has likewise  $0 \leq x - 1 \leq 1$ , so the right side of the two equations follows from the original functional equations. It is now a simple calculation to prove that all future extensions also satisfy

$$\begin{aligned} F(bx + a) - F(bx) &= F(b(x - 1) + a) - F(b(x - 1)) \\ &= F(a(x - 1) + b) - F(a(x - 1)) \\ &= F(ax + b) - F(ax), \end{aligned}$$

and similarly for  $f$ . The equations may be extended equally for  $x \rightarrow -\infty$ .

The extension process gains new meaning when we notice that the *differences*  $F(bx + a) - F(bx)$ ,  $b(f(bx + a) - f(bx))$ , etc. are *periodic*. Therefore, if  $F$  or  $f$  is of class  $L^2(0, 1)$  we may write, for all  $x$ ,

$$(E') \quad F(bx + a) - F(bx) = \sum \Gamma_n \exp(nx) = F(ax + b) - F(ax),$$

$$(e') \quad b(f(bx + a) - f(bx)) = \sum \gamma_n \exp(nx) = a(f(ax + b) - f(ax)),$$

where  $\exp(nx) = e^{2\pi inx}$  and  $\sum |\Gamma_n|^2 < \infty$ ,  $\sum |\gamma_n|^2 < \infty$ .

The extension process shows incidentally that the extended solutions will be locally in  $L^2$  for all  $x$ . The assumption that  $F$  be continuous on  $[0, 1]$  is not germane here. The extension procedure is still valid.

The work of the next section builds on the periodicity of the representation (e') to give all solutions of (e) for rational  $a$  and  $b$ .

**2. The rational case.** We begin with a special case, namely  $a = 1/n$  and  $b = (n-1)/n$  with  $n \geq 3$ . This problem was first undertaken by J. Dhombres [1] who provided a striking solution to (e) in the event that  $f$  was Riemann integrable. The fact that we will be able to extend his result should not overshadow the importance of Dhombres' work. He showed that the answer was the anticipated one for the class of functions one normally encounters in practice. Equally important is the fact that the Lebesgue integral could *not* be substituted in the argument by just changing a few words. Examples such as Dhombres' should be noted for both their creative as well as pedagogical value.

Assume now that  $f$  is a solution of (e). Multiply by  $\exp(-kx)$ , rearrange the terms and integrate:

$$a \int_0^1 (f(ax + b) - f(ax)) \exp(-kx) dx = b \int_0^1 (f(bx + a) - f(bx)) \exp(-kx) dx$$

or

$$\int_b^1 f(s) \exp\left(-k\left(\frac{s-b}{a}\right)\right) ds - \int_0^a f(s) \exp\left(-k\frac{s}{a}\right) ds$$

$$= \int_a^1 f(s) \exp\left(-k\left(\frac{s-a}{b}\right)\right) ds - \int_0^b f(s) \exp\left(-k\frac{s}{b}\right) ds.$$

Substitute the values  $a = 1/n$  and  $b = (n - 1)/n$  to give

$$\int_{1-1/n}^1 f(s) \exp(-kns) ds - \int_0^{1/n} f(s) \exp(-kns) ds$$

$$= \omega^k \int_{1/n}^1 f(s) \exp\left(-k\frac{n}{n-1}s\right) ds - \int_0^{1-1/n} f(s) \exp\left(-k\frac{n}{n-1}s\right) ds$$

where  $\omega = \exp(1/(n - 1))$ . When  $k = 0$  we gain no information. Otherwise we will write the preceding equation symbolically as

$$(4) \quad I_k - J_k = \omega^k K_k - L_k, \quad k = \pm 1, \pm 2, \dots$$

Then we can chart a chain of relations as follows.

$$I_1 - J_1 = \omega K_1 - L_1 = \omega^{n-1} K_{n-1} - L_{n-1}.$$

In addition to  $\omega^{n-1} = 1$  note that the integrals  $K_{n-1}$  and  $L_{n-1}$  produce the same values over the common interval of integration,  $[1/n, (n - 1)/n]$ , canceling one another. Once this simple relation is noticed, we press forward repeating the argument:

$$I_1 - J_1 = \omega K_1 - L_1 = \omega^{n-1} K_{n-1} - L_{n-1},$$

$$\omega^{n-1} K_{n-1} - L_{n-1} = I_{n-1} - J_{n-1} = \omega^{(n-1)^2} K_{(n-1)^2} - L_{(n-1)^2},$$

$$\omega^{(n-1)^2} K_{(n-1)^2} - L_{(n-1)^2} = I_{(n-1)^2} - J_{(n-1)^2}$$

$$= \omega^{(n-1)^3} K_{(n-1)^3} - L_{(n-1)^3} = \dots$$

From the Riemann-Lebesgue Lemma we conclude that all the differences involved vanish:

$$I_{(n-1)^k} - J_{(n-1)^k} = \omega^{(n-1)^k} K_{(n-1)^k} - L_{(n-1)^k} = 0, \quad k = 0, 1, 2, \dots$$

The same reasoning applies to all subsequent sequences not already part of an earlier chain, e.g.

$$I_2 - J_2 = \omega^2 K_2 - L_2$$

$$\vdots \quad \vdots \quad \vdots$$

$$I_{2(n-1)^2} - J_{2(n-1)^2} = \omega^{2(n-1)^2} K_{2(n-1)^2} - L_{2(n-1)^2}.$$

Therefore all of the differences on either side of (4) vanish. Note that the argument works equally well for negative indices.

The functions  $\{\exp(kns)\}$  form an orthogonal basis for  $L^2(0, 1/n)$  and  $L^2(1 - 1/n, 1)$ . To say that  $I_k - J_k = 0$  for  $k = \pm 1, \pm 2, \dots$  is the same as saying

that  $f$  restricted to  $[0, 1/n)$  is within a constant of  $f$  restricted to  $[1 - 1/n, 1)$ . That is,  $f(s + 1 - 1/n) = f(s) + c_1$  for  $0 \leq s < 1/n$ . The functions

$$\left\{ \exp\left(k \frac{n}{n-1} s\right) \right\} \quad \text{and} \quad \left\{ \omega^k \exp\left(k \frac{n}{n-1} s\right) \right\}$$

also form bases for  $L^2(0, 1 - 1/n)$  and  $L^2(1/n, 1)$  respectively. We conclude likewise that  $f(s + 1/n) = f(s) + c_2$  when  $0 \leq s < 1 - 1/n$ . Evidently  $f$  is a step-like function obtained by defining  $f$  arbitrarily on  $[0, 1/n)$  (square-integrable). On the interval  $[1/n, 2/n)$  the graph of  $f$  is the same as that of  $f$  on  $[0, 1/n)$  except that it is increased (decreased) by  $c_2$  units, and so on. Another way to view this is by defining  $s(x)$  to be the step function equal to the constant  $kc_2$  on  $[k/n, (k + 1)/n)$ ,  $k = 0, 1, \dots, n - 1$ . Set  $p(x)$  equal to  $f(x)$  on  $[0, 1/n)$  and extend  $p$  to be periodic of period  $1/n$  on  $[0, 1]$ . Then our solution becomes  $f(x) = s(x) + p(x)$ . It is easy to check that every such step function  $s$  is a solution of (e) as is every periodic function of period  $1/n$ .

If we take for a particular case  $p(x) = x$  on  $[0, 1/n)$  as our periodic function, then it is clear that the identity function  $j(x) = x$  can be expressed as  $j(x) = p(x) + s(x)$  where  $s$  has jumps of size  $1/n$ . So any step function is a linear combination of the identity and a periodic function.

**THEOREM 1.** *Every  $L^2$ -solution of (e) can be expressed as*

$$f(x) = \alpha x + p(x),$$

where  $p$  has period  $1/n$  and  $\alpha$  is a constant.

**COMMENT.** The reader is cautioned not to extend the validity of this extension of Dhombres' result beyond the case where  $a = 1/n$  and  $b = (n - 1)/n$ . However, it is possible that the theorem may be improved to include functions in, say,  $L^1(0, 1)$ . Our primary objective is to locate the extreme points of solutions  $f$  of (e) restricted to  $0 \leq f \leq 1$ .

**COROLLARY.** *If  $a = 1/n$  and  $b = (n - 1)/n$  then the extreme solutions of (e) within the class of functions satisfying  $0 \leq f \leq 1$  are*

- (1) *all characteristic functions of period  $1/n$ ;*
- (2) *the step function  $s$  given by  $s(x) = k/(n - 1)$ ,  $k/n \leq x < (k + 1)/n$ ,  $k = 0, 1, \dots, n - 1$ , and the function  $1 - s$ .*

Here is a sketch of how one might prove this assertion. Assume that  $f$  is a given solution with step size  $c > 0$  and that  $\delta < f < 1 - \delta$  on the set  $E \subset [0, 1/n)$  where  $\delta > 0$ . Choose

$$\varepsilon < \min(\delta, (n - 1)c).$$

Writing  $f(x) = p(x) + s(x)$  we note that

$$0 \leq p + (n - 1)c \leq 1$$

implies

$$0 \leq p + kc \leq 1 - (n - 1 - k)c$$

for  $k = 1, 2, \dots, n - 1$ . Set  $u(x) = \varepsilon \chi_E(x)$  on  $[0, 1/n)$  and assign it an increment of  $-\varepsilon/(n - 1)$  on the remaining intervals. Then for  $k/n \leq x < (k + 1)/n$

$$\begin{aligned} 0 &\leq f(x) + u(x) = p(x) + kc + \varepsilon(\chi_E(x) - k/(n - 1)) \\ &\leq 1 - (n - 1 - k)c + \varepsilon(1 - k/(n - 1)) \leq 1 \end{aligned}$$

and

$$0 \leq f(x) - u(x) = p(x) - \varepsilon\chi_E(x) + k(c - \varepsilon/(n - 1)) \leq 1.$$

Since  $0 \leq f \pm u \leq 1$ ,  $f$  cannot be extreme if the measure of  $E$  is positive. The remainder of the argument is left to the reader.

It is appropriate to check the corresponding extreme operators to see if any fail to preserve characteristic functions. In fact, none of them exhibits this property. To see this, let  $T$  be given by (1) where  $g$  is extreme in the class of solutions of (e) bounded between 0 and 1. If  $g$  is a characteristic function, then clearly  $T$  always carries characteristic functions into characteristic functions. As an example of how one handles the other extreme operators, consider the case where  $n = 3$ . Set

$$g(x) = \begin{cases} 0, & 0 \leq x < \frac{1}{3}, \\ \frac{1}{2}, & \frac{1}{3} \leq x < \frac{2}{3}, \\ 1, & \frac{2}{3} \leq x < 1. \end{cases}$$

Recall that  $\phi$  and  $\psi$  are defined by

$$\phi(x) = \begin{cases} 3x, & 0 \leq x < \frac{1}{3}, \\ \frac{3}{2}(x - \frac{1}{3}), & \frac{1}{3} \leq x \leq 1, \end{cases} \quad \psi(x) = \begin{cases} \frac{3}{2}x, & 0 \leq x < \frac{2}{3}, \\ 3(x - \frac{2}{3}), & \frac{2}{3} \leq x \leq 1. \end{cases}$$

Choose a set  $E_1 \subset [0, \frac{1}{2}]$  having positive measure less than  $\frac{1}{2}$ . Set  $E = E_1 \cup (\frac{1}{2} + E_1)$ , and let  $F = \phi^{-1}(E) \cap [0, \frac{2}{3}] \cup \psi^{-1}(E) \cap [\frac{2}{3}, 1]$ . A straightforward calculation verifies that  $T\chi_E = \chi_F$ .

At this point we return to the general case where  $a$  and  $b$  are rational. A classification of all  $L^2$ -solutions will be developed from an analysis of equation (e') and a class of fundamental solutions which were discovered earlier [6, 7]. We shall refer to them as  $\phi_n$ ,  $n = \pm 1, \pm 2, \dots$ . The definition is

$$\phi_n(x) = a(1 - \exp(n/a)) \exp(nx/b) + b(1 - \exp(n/b)) \exp(nx/a).$$

These functions are always solutions of (e) regardless of the nature of  $a$  and  $b$ . In the irrational case, the  $\phi_n$  appear as almost periodic functions, whereas for rational  $a$  and  $b$ , one obtains functions with periods of various lengths. In addition, one or both of the coefficients may vanish in the rational case.

An important distinction between the two cases is that the  $L^2(0, 1)$ -norms of the  $\phi_n$  are bounded away from zero when  $a$  and  $b$  are rational and  $\phi_n$  does not vanish identically. We could, therefore, assume that  $\|\phi_n\|_2 = 1$  without causing any conflict later.

If  $f$  is a solution of (e) in  $L^2(0, 1)$  we will assume that an extension has been made to  $\mathbf{R}$  such that (e') obtains

$$b(f(bx + a) - f(bx)) = \sum \gamma_\nu \exp(\nu x) = a(f(ax + b) - f(ax))$$

with  $\sum |\gamma_\nu|^2 < \infty$ . Looking at these differences for  $\phi_n$  we obtain

$$a(\phi_n(ax + b) - \phi_n(ax)) = ab(1 - \exp(n/b))(\exp(n/a) - 1) \exp(nx).$$

For those  $m$  with

$$(1 - \exp(m/b))(\exp(m/a) - 1) \neq 0$$

we form the difference

$$g = f - \sum \tilde{\gamma}_m \phi_m.$$

where

$$\tilde{\gamma}_m = \gamma_m [ab(1 - \exp(m/b))(\exp(m/a) - 1)]^{-1}.$$

Then  $g$  is again an  $L^2$ -solution of (e) and

$$(5) \quad b(g(bx + a) - g(bx)) = \sum_{\nu \in N'} \gamma_\nu \exp(\nu x) = a(g(ax + b) - g(ax)),$$

where  $N'$  is the set of integers  $n'$  such that  $n'/a$  or  $n'/b$  is an integer. We can see what these values must be. Set  $a = \alpha/(\alpha + \beta)$ ,  $b = \beta/(\alpha + \beta)$ , where  $\alpha$  and  $\beta$  are natural numbers and  $(\alpha, \beta) = 1$ . Then  $N' = \{\pm\alpha\mathbf{N} \cup \pm\beta\mathbf{N}\}$ . If, for instance,  $a = 3/10$  and  $b = 7/10$  then  $N' = \{0, \pm 3, \pm 6, \pm 7, \dots, \pm 21, \dots\}$ . It is worth emphasizing that for  $\nu \in N'$  it does no good to subtract  $\tilde{\gamma}_\nu \phi_\nu$  from  $f$ , since  $\phi_\nu$  will be periodic of period  $(\alpha + \beta)^{-1}$  and no information is supplied to equation (5). Nevertheless, there may be terms in the expansion with  $\nu \in N'$  which do not come from periodic functions—that is the most interesting part of this analysis.

Choose integers  $p$  and  $q$  such that  $p\beta - q\alpha = 1$ . Then by changing variables and bearing in mind that (5) is valid for all real values of  $x$ , we can write

$$g(x + (q - l + 1)a) - g(x + (q - l)a) = b^{-1} \sum \gamma_\nu \exp\left(\nu \frac{x}{b}\right) \omega_b^{\nu(q-l)},$$

$$g(x + (p - k + 1)b) - g(x + (p - k)b) = a^{-1} \sum \gamma_\nu \exp\left(\nu \frac{x}{a}\right) \omega_a^{\nu(p-k)},$$

where  $1 \leq l \leq q$ ,  $1 \leq k \leq p$ , and  $\omega_a = \exp(a^{-1})$ ,  $\omega_b = \exp(b^{-1})$ . If we sum on  $k$  and  $l$  the left sides collapse to

$$g(x + qa) - g(x) = b^{-1} \sum \gamma_\nu \exp\left(\nu \frac{x}{b}\right) d_\nu,$$

$$g(x + pb) - g(x) = a^{-1} \sum \gamma_\nu \exp\left(\nu \frac{x}{a}\right) c_\nu,$$

with

$$d_\nu = \begin{cases} (\omega_b^{q\nu} - 1)[(\omega_b^\nu - 1)]^{-1} & \text{if } \beta \nmid \nu, \\ q & \text{if } \beta \mid \nu, \end{cases}$$

$$c_\nu = \begin{cases} (\omega_a^{p\nu} - 1)[(\omega_a^\nu - 1)]^{-1} & \text{if } \alpha \nmid \nu, \\ p & \text{if } \alpha \mid \nu. \end{cases}$$

Subtracting and substituting again, we obtain

$$(6) \quad g(x + pb - qa) = g(x + (\alpha + \beta)^{-1})$$

$$= g(x) + \sum_{\gamma \in N'} \left[ a^{-1} \gamma_\nu \exp\left(\nu \frac{x}{a}\right) c_\nu - b^{-1} \gamma_\nu \exp\left(\nu \frac{x}{b}\right) \bar{\omega}_b^{\nu q} d_\nu \right].$$

We must show that this representation is independent of our choice of  $p$  and  $q$  satisfying  $p\beta - q\alpha = 1$ . It is easy to see that if  $p'\beta - q'\alpha = 1$  then  $p' = p + k\alpha$  and  $q' = q + k\beta$  for some integer  $k$ . Therefore terms such as  $(\omega_a^{p'\nu} - 1)$  would equal  $(\omega_a^{p\nu} - 1)$ . But we are mainly concerned with those  $\nu \in N'$  where, in each term of the series (6),  $c_\nu = p$  and  $d_\nu = q$ . To see that this dependence on  $p$  and  $q$  is only apparent, note that if  $\nu = \alpha k$  for some  $k$ , then  $\exp(\nu x/a) = \exp(k(\alpha + \beta)x)$ . This exponential occurs in exactly one other place in the series, namely when  $\nu' = \beta k$

for this same  $k$  ( $\bar{\omega}_b^{\nu'q} = 1$  as well). Multiply (5) by  $\exp(-\nu x)$  and integrate. The right side can be arranged to yield

$$\gamma_\nu = \int_b^1 - \int_0^a g(x) \exp(-k(\alpha + \beta)x) dx.$$

Multiplying the left side of (5) by  $\exp(-\nu'x)$  we obtain

$$\gamma_{\nu'} = \int_a^1 - \int_0^b g(x) \exp(-k(\alpha + \beta)x) dx.$$

These integrals are clearly the same. Examination of (6) will confirm that the coefficient of  $\exp(k(\alpha + \beta)x)$  is  $\gamma_\nu(pb - qa)(ab)^{-1}$  which is independent of the choice of  $p$  and  $q$ .

Writing  $(\alpha + \beta)^{-1} = d$ , we may view (6) as representing  $g$  as  $g(x+d) = g(x) + h(x)$ . Any periodic function of period  $d$  will always be a solution of (e) and may be added to or subtracted from  $g$ . Some, but not all, of the terms in the series for  $h$  are also of period  $d$ , as we have seen. By choosing a  $d$ -periodic function equal to  $g$  on  $[0, d]$  and subtracting (for the moment) we may assume that  $g = 0$  on this interval. Then we have for  $0 \leq x \leq d$ ,

$$\begin{aligned} g(x) &= 0 \\ g(x + d) &= h(x) \\ &\vdots \quad \vdots \quad \vdots \\ g(x + kd) &= h(x) + h(x + d) + \dots + h(x + (k - 1)d) \end{aligned}$$

The graph of  $g$  is built as a series of steps along intervals of length  $d$ .

Equation (6) gives a necessary condition on  $g$  which is not sufficient for a solution of (e). There are further restrictions on the coefficients  $\gamma_\nu$  which remain to be determined. They involve interesting arithmetic relations linking  $\alpha$  and  $\beta$ . To see what these must be, we multiply (5) by  $\exp(-\nu x)$  for  $\nu \in N'$  and integrate, writing  $a$  and  $b$  in terms of  $\alpha$  and  $\beta$ :

$$\begin{aligned} \omega_b^\nu \int_a^1 - \int_0^b g(x) \exp\left(-\nu(\alpha + \beta)\frac{x}{\beta}\right) dx \\ = \gamma_\nu = \omega_a^\nu \int_0^1 - \int_0^a g(x) \exp\left(-\nu(\alpha + \beta)\frac{x}{\alpha}\right) dx. \end{aligned}$$

Because of the choice of  $\nu$  at least one exponent will always be integral. Working with examples, one can see what the relationships between the  $\gamma_\nu$  must be. For instance, let  $\alpha = 2$  and  $\beta = 5$  ( $a = \frac{2}{7}$ ,  $b = \frac{5}{7}$ ). Then we calculate that

$$\begin{aligned} \gamma_2 &= \gamma_5 \\ \gamma_4 &= \gamma_{10} = \gamma_{25} \\ \gamma_6 &= \gamma_{15} \\ \gamma_8 &= \gamma_{20} = \gamma_{50} = \gamma_{125} \\ &\dots \end{aligned}$$

To discover the pattern, it is useful to rewrite the sequence of indices as follows:

$$\begin{array}{ccccccc}
 & & 2 & & 5 & & \\
 & & 2^2 & & 2 \cdot 5 & & 5^2 \\
 & 2^3 & 2^2 \cdot 5 & & 2 \cdot 5^2 & & 5^3 \\
 & \dots & \dots & \dots & \dots & \dots & \dots \\
 & & 2 \cdot 3 & & 3 \cdot 5 & & \\
 & & 2^2 \cdot 3 & & 2 \cdot 3 \cdot 5 & & 3 \cdot 5^2 \\
 2^3 \cdot 3 & & 2^2 \cdot 3 \cdot 5 & & 2 \cdot 3 \cdot 5^2 & & 3 \cdot 5^3 \\
 & \dots & \dots & \dots & \dots & \dots & \dots \\
 & & 2 \cdot 7 & & 7 \cdot 5 & & 
 \end{array}$$

and so on. Attached to each row are two natural numbers, a number  $\mu \in N \sim N'$  and an exponent  $m > 0$ . To calculate the row of indexes for which the  $\gamma_\nu$  are equal simply write down

$$2^m \cdot \mu \quad 2^{m-1} \cdot \mu \cdot 5 \quad \dots \quad 2 \cdot \mu \cdot 5^{m-1} \quad \mu \cdot 5^m.$$

For the general case, one chooses any  $\mu \in N \sim N'$ ,  $m \in N$  and calculate that all the  $\gamma_\nu$  with  $\nu = \alpha^k \mu \beta^{m-k}$ ,  $1 \leq k \leq m$ , are equal. A complementary set is generated when  $\nu$  is replaced by  $-\nu$ . This is a straightforward computation which will not be carried out here. In checking this, the interested reader will note that the coefficients  $\omega_a$  and  $\omega_b$  are equal to 1 at the proper times.

The equality of groups of coefficients  $\gamma_\nu$  restricts the admissible series in the representation (6). In view of our identifications, it now appears appropriate to rewrite each coefficient as  $\gamma_{\mu,m}$  and group terms accordingly. The series in (6) then becomes

$$(7) \quad s(x) = \sum \gamma_{\mu,m} \sum_{k=0}^m \left[ a^{-1} \exp\left(\alpha^k \beta^{m-k} \mu \frac{x}{a}\right) c_k - b^{-1} \exp\left(\alpha^k \beta^{m-k} \mu \frac{x}{b}\right) d_k \right]$$

where

$$c_k = \begin{cases} p, & k \neq 0, \\ (\omega_a^{\beta^{m-1}} - 1)[(\omega_a^{\beta^m} - 1)]^{-1}, & k = 0, \end{cases}$$

$$d_k = \begin{cases} q, & k \neq m, \\ (\omega_b^{\alpha^{m-1}} - 1)[(\omega_b^{\alpha^m} - 1)]^{-1}, & k = m. \end{cases}$$

The number  $\mu$  can be any nonzero integer not divisible by  $\alpha$  or  $\beta$ , and  $m$  is a natural number. We condense these observations into the next theorem.

**THEOREM 2.** *Given any rational numbers  $a$  and  $b$  in  $(0, 1)$  satisfying  $a + b = 1$  with  $a = \alpha/(\alpha + \beta)$ ,  $b = \beta/(\alpha + \beta)$ ,  $(\alpha, \beta) = 1$ , the most general  $L^2$ -solution of equation (e) is obtained as follows. Define  $f$  arbitrarily on  $[0, (\alpha + \beta)^{-1}]$ . Construct an extension by setting*

$$f(x + (\alpha + \beta)^{-1}) = f(x) + s(x)$$

where  $s$  is given by (7). Add to  $f$  any  $L^2$ -periodic function of period  $(\alpha + \beta)^{-1}$ . The coefficients  $\gamma_{\mu,m}$  must be chosen so that

$$\sum (m + 1) |\gamma_{\mu,m}^2| < \infty$$

to ensure  $L^2$ -convergence of the series.

The function  $s$  which contributes to the step-like behavior of a solution to (e) is made up of two parts. One is a function of period  $a$  and the second of period  $b$ . Because of the coefficients involved they are not the same functions. In some simple cases ( $\alpha = 2$  and  $\beta = 3$  for instance) one can take  $p = q = 1$  and the series becomes much more manageable. Even in these cases it is still difficult to identify the extremal elements in the class of solutions bounded between 0 and 1. It is true that when  $p = q = 1$  we can write

$$s(x) = a^{-1}g(x/a) - b^{-1}g(x/b)$$

for some function  $g$  and get some graphic idea of what a solution must look like. The problem is that we cannot use *arbitrary*  $g$ . There are relations between the Fourier coefficients of the  $s$ -functions which must be maintained. Each admissible  $s$  is a sum of lacunary trigonometric polynomials whose behavior is hardly transparent.

The theorem is not proved until we demonstrate that all solutions can be obtained in the manner described. To that end we refer to the differences  $g = f - \sum \gamma_m \phi_m$ . Define

$$\sigma_{\mu,m}(x) = \sum \left[ a^{-1} \exp\left(\alpha^k \beta^{m-k} \mu \frac{x}{a}\right) c_k - b^{-1} \exp\left(\alpha^k \beta^{m-k} \mu \frac{x}{b}\right) d_k \right].$$

Then define  $s_{\mu,m}(x) = 0$  for  $0 \leq x < (\alpha + \beta)^{-1}$  and recursively on  $[0, 1]$  by

$$s_{\mu,m}(x + (\alpha + \beta)^{-1}) = s_{\mu,m}(x) + \sigma_{\mu,m}(x).$$

A tedious, but not very difficult computation will verify that

$$\begin{aligned} b[s_{\mu,m}(bx + a) - s_{\mu,m}(bx)] &= \sum \exp(\alpha^k \beta^{m-k} \mu x) \\ &= a[s_{\mu,m}(ax + b) - s_{\mu,m}(ax)]. \end{aligned}$$

There is no loss in assuming that  $s_{\mu,m}$  has been extended to  $\mathbf{R}$  as a solution of (e). If we subtract  $\sum \gamma_{\mu,m} s_{\mu,m}$  from  $g$  and set  $h = f - \sum \tilde{\gamma}_n \phi_n - \sum \gamma_{\mu,m} s_{\mu,m}$  we arrive at a solution  $h$  of (e) which satisfies

$$b(h(bx + a) - h(bx)) = \text{const.} = \gamma_0 = a(h(ax + b) - h(ax))$$

for  $0 \leq x \leq 1$ . Let  $\delta_0(x) = \gamma_0(\alpha + \beta)$  for  $0 \leq x < (\alpha + \beta)^{-1}$ , and

$$s_0(x + (\alpha + \beta)^{-1}) = s_0(x) + \delta_0(x),$$

its extension as a step function. Subtract this function from  $g$  as well. We will be left with an  $L^2$ -function which satisfies the simultaneous equations

$$h(bx + a) - h(bx) = 0, \quad h(ax + b) - h(ax) = 0.$$

Such a function must be periodic with period  $(\alpha + \beta)^{-1}$ . This is quickly verified by assuming that  $h$  has been extended to  $\mathbf{R}$  with the two equations still in force. A repetition of an earlier argument gives

$$h(x + qb) - h(x) = 0, \quad h(x + pa) - h(x) = 0,$$

and so

$$h(x + qb - pa) - h(x) = h(x + (\alpha + \beta)^{-1}) - h(x) = 0.$$

The most general form of an  $L^2$ -solution to (e) becomes, in our notation,

$$f(x) = h(x) + \sum \gamma_n \phi_n(x) + \sum \gamma_{\mu,m} s_{\mu,m}(x),$$

where  $s_0$  is absorbed in the third term.

One may view this as consisting of a purely periodic part, a “mixed” part consisting of functions with period  $\alpha\beta$  and a series of step-like functions. The third series should not be thought of as representing a simple monotonic function. The functions involved are periodic themselves and steps may alternately increase or decrease. This alone makes the identification of extreme points in the range  $0 \leq f \leq 1$  very difficult, aside from the obvious periodic characteristic functions.

One final comment should be made. In the construction of the functions  $s_{\mu,m}$  we always assumed that they were equal to 0 on  $[0, (\alpha + \beta)^{-1}]$ . This may give the appearance of an unjustified restriction. No loss actually occurs, however; we could set  $s_{\mu,m}$  equal to anything we wish on this interval. The loss or gain is offset by the periodic term  $h$ . So this initialization is merely a convenience which in no way affects the solution—only its form.

**3. The irrational case.** Results in this case are less complete than in the rational one. Indeed, determining even a dense set of solutions is beyond the scope of this work. We will investigate the ‘integrated’ equation (E) in this section. Essentially complete descriptions of solutions are given in [1] and [7] for the case where the solution is assumed to be bounded and almost-periodic on the real line or entire of exponential growth.

Recall that any solution  $F$  of (E), extended to  $\mathbf{R}$ , can be expressed by (E’):

$$F(bx + a) - F(bx) = \sum \Gamma_n \exp(nx) = F(ax + b) - F(ax).$$

When  $a$  and  $b$  are irrational the functions

$$\Phi_n(x) = (1 - \exp(n/a)) \exp(nx/b) + (1 - \exp(n/b)) \exp(nx/a)$$

satisfy

$$\begin{aligned} \Phi_n(bx + a) - \Phi_n(bx) &= (1 - \exp(n/a))(\exp(n/b) - 1) \exp(nx) \\ &= \Phi_n(ax + b) - \Phi_n(ax). \end{aligned}$$

The coefficient of  $\exp(nx)$  may be rewritten as

$$4\lambda_n \sin(\pi n/a) \sin(\pi n/b) = \delta_n,$$

where  $|\lambda_n| = 1$ . If we were to form the difference

$$h(x) = F(x) - \sum \Gamma_n \delta_n^{-1} \Phi_n(x)$$

a solution  $h$  of (E) would result with the property that

$$h(bx + a) - h(bx) = h(ax + b) - h(ax) = 0.$$

However, the coefficients  $\delta_n$  form small divisors—a problem which does not occur in the rational case. A linear term may be added to the series by adjoining the solution  $\Phi_0(x) = x^2 - x$  to the family  $\{\Phi_n\}$ . This leads one to suspect that the functions  $\{\Phi_n\}$  are only useful in representing almost-periodic or entire solutions of (E) and that other solutions may exist. First we establish an estimate of the growth of solutions of (E) on the unit interval.

PROPOSITION 1. *The linear operation which associates a function defined in  $[0, b)$  with its extension as a solution of (E) is  $L^2$ -bounded.*

PROOF. Assume that  $F \in L^2(0, b)$  is arbitrary and let  $\lambda = a/b$ . For any  $x$  satisfying  $0 < x < 1 - \lambda$ , we have by [7, equation (4)]

$$F(1 - a\lambda^n(1 - x)) = F(bx + a) + \sum_{k=0}^n [F(a(1 - \lambda^k(1 - x))) - F(b(1 - \lambda^k(1 - x)))]$$

where the symbol also refers to the extended function  $F$ . Changing variables, we obtain for the extension

$$(8) \quad F(s) = F(1 - \lambda^{-(n+1)}(1 - s)) + \sum_{k=0}^n [F(a - \lambda^{-k}(1 - s)) - F(b - \lambda^{-(k+1)}(1 - s))],$$

valid for  $1 - a\lambda^n \leq s < 1 - a\lambda^{n+1}$  and  $n = 0, 1, 2, \dots$ . The function  $G$  given by  $G(s) = F(1 - \lambda^{-(n+1)}(1 - s))$  for  $1 - a\lambda^n \leq s < 1 - a\lambda^{(n+1)}$  is easily seen to belong to  $L^2(b, 1)$ . For

$$\int_b^{1-a\lambda^{(n+1)}} |G(s)|^2 ds = \sum_{k=0}^n \int_{1-a\lambda^k}^{1-a\lambda^{(k+1)}} |F(1 - \lambda^{-(k+1)}(1 - s))|^2 ds \leq \sum_{k=0}^\infty \lambda^{(k+1)} \int_a^b |F(s)|^2 ds = \frac{\lambda}{1-\lambda} \int_a^b |F(s)|^2 ds.$$

It is less obvious, but nevertheless true, that the function  $J$  defined by

$$J(s) = \sum_{k=0}^n F(a - \lambda^{-k}(1 - s)),$$

$1 - a\lambda^n \leq s < 1 - a\lambda^{(n+1)}$ , also lies in  $L^2(b, 1)$ , for

$$|J(s)|^2 \leq \sum_{k=0}^n (k + 1)^2 |F(a - \lambda^{-k}(1 - s))|^2 \sum_{k=0}^n (k + 1)^{-2}.$$

Integration yields

$$\int_{1-a\lambda^n}^{1-a\lambda^{(n+1)}} |J(s)|^2 ds \leq \frac{\pi^2}{6} \sum_{k=0}^n (k + 1)^2 \lambda^k \int_{a(1-\lambda^{(n-k)})}^{a(1-\lambda^{(n-k+1)})} |F(s)|^2 ds.$$

Set

$$I_m = \frac{\pi^2}{6} \int_{a(1-\lambda^m)}^{a(1-\lambda^{(m+1)})} |F(s)|^2 ds, \quad m = 0, 1, 2, \dots$$

Then

$$\int_b^{1-a\lambda^{(n+1)}} |J(s)|^2 ds \leq I_0 + 2^2 \lambda I_0 + I_1 + 3^2 \lambda^2 I_0 + 2^2 \lambda I_1 + I_2 + \dots + (n + 1)^2 \lambda^n I_0 + n^2 \lambda^{n-1} I_1 + \dots + I_n.$$

Evidently

$$\int_b^1 |J(s)|^2 ds \leq \sum_{k=0}^{\infty} (k+1)^2 \lambda^k \sum_{m=0}^{\infty} I_m.$$

Similarly, one shows that the function  $K$  defined by

$$K(s) = \sum F(b - \lambda^{-(k+1)}(1 - s))$$

for  $1 - a\lambda^n \leq s < 1 - a\lambda^{(n+1)}$  satisfies an inequality of the type

$$\int_b^1 |K(s)|^2 ds \leq C \int_0^b |F(s)|^2 ds,$$

completing the argument.

One of the reasons for investigating equation (E) is that this proposition fails when one considers extensions of functions which are required to satisfy equation (e). For the sake of illustration, assume that  $a < \frac{1}{3}$ . Define  $f = 0$  on  $[0, a)$  and  $f = 1$  on  $[a, b)$ . The extension of  $f$  to  $[0, 1)$  as a solution of (e) satisfies the following with  $A$  and  $B$  positive:

$$f(s) \geq A\lambda^{-(n+1)} + B, \quad 1 - a\lambda^n \leq s < 1 - a\lambda^{(n+1)},$$

for  $n \geq 1$ . A direct calculation will bear this out. Integration takes place on intervals whose length is proportional to  $\lambda^n$ . We see that this very simple function does not extend to a function even in  $L^1(b, 1)$ .

The extension process whereby functions defined on  $[0, b)$  are extended to solutions of (E) on  $[0, 1)$  is a linear one. This observation, used in conjunction with the last proposition, tells us that functions in a dense subset of  $L^2(0, b)$  extend to a dense set of solutions of (E) on the unit interval. If the functions  $\{\Phi_n\}$  are to play a prominent role in the investigation of solutions of (E), one should expect their linear span to be dense. We must first add to this class one additional solution,  $\Phi_{00}(x) = 1$ . In general, the set  $\{\Phi_n\}$ ,  $n = 00, 0, \pm 1, \pm 2, \dots$ , may not span an  $L^2$ -dense subspace of solutions of (E). In certain cases, one cannot find nontrivial functions orthogonal to the class  $\{\Phi_n\}$  on the interval  $[0, b)$ . The question as to whether or not this is true for all choices of irrational values of  $a$  and  $b$  remains unsettled. We shall work with specific choices of these numbers, chosen so that  $\exp(n/a) = \exp(n/b)$ ,  $n = \pm 1, \pm 2, \dots$ . In other words,  $a^{-1} = b^{-1} + N$  for a natural number  $N$ . The simplest case occurs when  $N = 1$  giving  $a = \frac{1}{2}(3 - \sqrt{5})$  and  $b = \frac{1}{2}(\sqrt{5} - 1)$ . Assume that a particular value of  $N$  has been chosen. Then any function  $f$  orthogonal to the class  $\{\Phi_n\}$  on  $[0, b)$  must satisfy

$$(9) \quad \int_0^b \exp\left(\frac{n}{b}x\right) f(x) dx = - \int_0^b \exp\left(\frac{n}{a}x\right) f(x) dx,$$

$n = 0, \pm 1, \pm 2, \dots$ . Dividing the integrals on the left by  $\sqrt{b}$  produces the set of Fourier coefficients of  $f$  on  $[0, b)$ . The orthogonality with  $\Phi_0$  will be taken up later. Write

$$f = f\chi_1 + f\chi_2 + \dots + f\chi_N f\chi_{N+1} = f_1 + f_2 + \dots + f_{N+1}$$

where  $\chi_\nu$  is the characteristic function of the interval  $[(\nu - 1)a, \nu a)$ ,  $\nu = 1, 2, \dots, N + 1$ . One verifies that  $Na < b < (N + 1)a$ , so that the meaning of the last term



Equivalently, we may transfer all the information to  $[0, a)$  by writing

$$f_k(x) = -\lambda \sum f_\nu(\lambda x + \lambda ka), \quad 0 \leq x < a,$$

$k = 2, 3, \dots, N$ . Note that all the values  $\lambda x + \lambda ka$  lie in the interval  $[0, a)$ . Defining

$$u(x) = \sum f_\nu(x), \quad 0 \leq x < a,$$

we convert the system of functional equations to a single one by adding:

$$(14) \quad u(x) = -\lambda(u(\lambda x) + u(\lambda x + \lambda a) + \dots + u(\lambda x + \lambda(N - 1)a)) + u(\lambda^2 x + \lambda Na).$$

We may recover the functions  $f_\nu$  from a solution of (14) and therefore a solution  $f$ . Moreover, an  $L^2$ -solution of (14) produces  $L^2$ -solutions of (13) and vice versa.

Solutions of (14) in  $L^p(0, a)$ ,  $0 < p < 1$ , are easily obtained. If  $0 < x < \lambda Na$ , it is also true that  $0 < \lambda x + \lambda ka < \lambda Na$ ,  $k = 1, 2, \dots, N - 1$ . On rewriting (14) as

$$(15) \quad u(\lambda^2 x + \lambda Na) = u(x) + \lambda \sum_{k=0}^{N-1} u(\lambda x + \lambda ka),$$

we see that when  $u$  is given for  $0 \leq x < \lambda Na$  it is also defined as a solution of (15) for  $\lambda Na < \lambda^2 x + \lambda Na < \lambda Na + \lambda^3 Na$ . The process of extension from  $\lambda Na + \lambda^3 Na$  to  $\lambda Na + \lambda^3 Na + \lambda^5 Na$ , etc. should be familiar by now. Ultimately we may extend  $u$  as a solution to

$$a = \lambda Na + \lambda^3 Na + \lambda^5 Na + \dots$$

(observing that the choice of  $a, b$  and  $N$  requires that  $\lambda$  satisfy  $\lambda^2 + \lambda N - 1 = 0$ ). If  $u \in L^p(0, \lambda Na)$  then simple estimates show that the extension, denoted by  $u_e$ , satisfies

$$\|u_e\| \leq C_p \|u\|_p$$

where  $\|u\|_p$  is the  $L^p(0, \lambda Na)$  norm of  $u$ . Although there is always a function which uniquely extends a given  $u$ , one cannot say whether it will be square-integrable, since the constants  $C_p \rightarrow \infty$  as  $p \rightarrow 1^-$ . And for good reason: If  $u(x) = 1$  for  $0 \leq x < \lambda Na$ , then using (15) we may calculate directly that the extension  $u_e$  fails to belong to  $L^1(0, a)$ . The linear mapping  $u \rightarrow u_e$  is unbounded from  $L^1(0, \lambda Na)$  to  $L^1(0, a)$ . Since the extension process is monotone, it is not difficult to see that any nonnegative  $u \in L^1(0, \lambda Na)$  fails to extend with finite  $L^1$ -norm ( $u \neq 0$ ). Several other properties relating to the extension may be derived, yet the actual construction of an  $L^1$ -solution of (14) or (15) different from 0 is seen to be impossible. In what follows, it will suffice to assume that all functions are real-valued.

**THEOREM 3.** *There are no solutions of (15) in  $L^1(0, a)$  except for the null function.*

**COROLLARY.** *The class  $\{\Phi_n\}$ ,  $n = 00, \pm 1, \pm 2, \dots$  spans  $L^1(0, b)$ , hence spans the class of  $L^2$ -solutions of (E) on  $(0, 1)$ .*

The proof of the theorem uses some ideas from ergodic theory. If  $u \in L^1(0, a)$  were a solution of (15), then on taking absolute values, we should have

$$\begin{aligned} \lambda^2 |u(\lambda^2 x + \lambda Na)| &\geq |u(x)| - \lambda \left| \sum u(\lambda x + \lambda ka) \right| \\ &\geq |u(x)| - \lambda \sum |u(\lambda x + \lambda ka)|. \end{aligned}$$

Integrating both sides, we obtain

$$\begin{aligned} \lambda^2 \int_0^a |u(\lambda^2 x + \lambda Na)| dx &= \int_{\lambda Na}^a |u(x)| dx \\ &\geq \int_0^a |u(x)| dx - \lambda \sum \int_{\lambda ka}^{\lambda(k+1)a} |u(x)| dx. \end{aligned}$$

But the right side reduces to the left. Equality holds following use of the triangle inequality (twice), from which we conclude that  $u(x)$  and each term  $u(\lambda x + \lambda ka)$  have opposite signs for almost all  $x$  in  $(0, a)$ . That is,

$$\lambda^2 |u(\lambda^2 x + Na)| = |u(x)| = \lambda \sum |u(\lambda x + \lambda ka)|$$

(so  $u(\lambda^2 x + \lambda Na)$  and  $u(x)$  will have the same sign or vanish together). Rewrite this equation as

$$|u(x)| = \lambda \sum |u(\lambda x + \lambda ka)| + \lambda^2 |u(\lambda^2 x + \lambda Na)|.$$

Note that the transformation

$$(Tv)(x) = \lambda \sum v(\lambda x + \lambda ka) + \lambda^2 v(\lambda^2 x + \lambda Na)$$

satisfies conditions (i), (ii), and (iii) at the beginning of this article (except that  $L^1(0, 1)$  is replaced by  $L^1(0, a)$ ). In fact,  $T$  is seen to be the adjoint of the operator

$$(T^*w)(x) = \sum w\left(\frac{x - \lambda ka}{\lambda}\right) \chi_{k+1}(x) + w\left(\frac{x - \lambda Na}{\lambda^2}\right) \chi_{N+1}(x),$$

where the functions  $\chi_1, \dots, \chi_N$  are the same as before, while  $\chi_{N+1}$  is the characteristic function of  $[\lambda Na, a)$ . One should view  $T^*$  as the composition operator  $T^*w = w \circ \rho$ , where

$$\rho(x) = \begin{cases} (x/\lambda) \pmod{a}, & 0 \leq x < \lambda Na, \\ (x - \lambda Na)/\lambda^2, & \lambda Na \leq x < a. \end{cases}$$

It is not difficult to show that  $\rho$  is measure preserving on  $(0, a)$ . It follows from [3] or [10] that  $\rho$  is ergodic as well. This can be seen directly by performing a ‘gedanken’ experiment: the inverse images of the intervals  $(0, \lambda a)$ ,  $(\lambda a, 2\lambda a)$ ,  $\dots$ ,  $((N - 1)\lambda a, \lambda Na)$ ,  $(\lambda Na, a)$  form a base for the open sets of  $(0, a)$  in the sense that any open subset of  $(0, a)$  can be expressed as a countable union of the inverse images except possibly for a denumerable set. If  $E$  is a set of positive measure which is invariant under the action of  $\rho$ , then by analyzing the inverse images of  $E$  it becomes clear that  $E$  intersects each open interval in a set of the same proportionality. There exists a  $\theta > 0$  such that  $|E \cap (c, d)| \geq \theta|d - c|$  so that  $E$  must have full measure equal to  $a$ .

Linear transformations which satisfy conditions (i), (ii), and (iii) on  $L^1$  have adjoints defined on  $L^\infty$ . However, the adjoints share the same three properties and have unique extensions to  $L^1$  with no increase in norm. Details demonstrating this may be found in [4] although a direct verification is straightforward. Starting from  $T|u| = |u|$  we set  $E_s = \{|u| \geq s\}$  and let  $\chi_{E_s}$  be the characteristic function of  $E_s$ . Then, using inner product notation we arrive at

$$\langle |u|, \chi_{E_s} \rangle = \langle T|u|, \chi_{E_s} \rangle = \langle |u|, \chi_{E_s}(\rho) \rangle = \langle |u|, \chi_{\rho^{-1}(E_s)} \rangle.$$

This could only occur if  $\rho^{-1}(E_s) = E_s$  (a.e.). Evidently,  $T^*|u| = |u|$  as well. As  $\rho$  is ergodic,  $|u|$  must be constant. We want to establish that this constant must be 0. Assuming otherwise, let  $|u(x)| = 1$  on  $(0, a)$ , where  $u$  is a solution of (15). The interval  $[0, \lambda Na)$  is composed of two sets,  $Q^+ = \{u(x) = 1\}$  and  $Q^- = \{u(x) = -1\}$ . To obtain the remaining information about  $u$  we recall our observation that  $u(\lambda^2 x + \lambda Na)$  must be 1 when  $u(x) = 1$  and vice versa. Therefore we may develop the remainder of the set where  $u(x) = 1$  by recursion:

$$\begin{aligned} \{u(x) = 1\} &= Q^+ \cup \{\lambda^2 Q^+ + \lambda Na\} \cup \{\lambda^4 Q^+ + \lambda^3 Na + \lambda Na\} \cup \dots, \\ \{u(x) = -1\} &= Q^- \cup \{\lambda^2 Q^- + \lambda Na\} \cup \{\lambda^4 Q^- + \lambda^3 Na + \lambda Na\} \cup \dots. \end{aligned}$$

Suppose  $u(x) = 1$ . Then  $u(y) = -1$  for  $y = \lambda x, \lambda x + \lambda a, \dots, \lambda x + \lambda Na$ . In addition, by what we have just observed,  $u(y) = -1$  also for  $y = \lambda^3 x + \lambda Na, \lambda^5 x + \lambda^3 Na + \lambda Na, \dots, \lambda^3 x + \lambda^3 a + \lambda Na, \lambda^5 x + \lambda^5 a + \lambda^3 Na + \lambda Na, \dots$ , etc. Perhaps a simpler way to look at this would be to define a function

$$\alpha(x) = \begin{cases} x/\lambda, & 0 \leq x < \lambda Na, \\ (x - \lambda Na)/\lambda^3, & \lambda Na \leq x < \lambda Na + \lambda^3 Na, \\ (x - \lambda Na - \lambda^3 Na)/\lambda^5, & \lambda Na + \lambda^3 Na \leq x < \lambda Na + \lambda^3 Na + \lambda^5 Na, \\ \vdots & \vdots \quad \vdots \quad \vdots \end{cases}$$

etc. This will define  $\alpha$  on  $[0, a)$ . Take

$$\beta(x) = \alpha(x) \pmod{a}.$$

Then  $\beta^{-1}(Q^+) = Q^-$ ,  $\beta^{-1}(Q^-) = Q^+$ , and  $\beta$  is measure preserving. The complete symmetry between  $Q^+$  and  $Q^-$  dictates that for any interval  $(c, d) \subset (0, a)$ ,  $|Q^+ \cap (c, d)| = |Q^- \cap (c, d)| = \frac{1}{2}|d - c|$ . The two sets  $Q^+$  and  $Q^-$  must have measure equal to  $a$ , which is impossible. Hence  $u$  must be null and (15) has only trivial solutions in  $L^1(0, a)$ .

A consequence of this result is that the functions  $\Phi_n$  are dense in  $L^2(0, b)$  without the inclusion of the quadratic  $\Phi_0$ . It is difficult at this point to understand its role in the analysis of equation (E), although this same function plays a fundamental one in the study of entire solutions of (E) which are of exponential type (see [7]). Also unclear is how one might go over to the general case where  $a$  and  $b$  are arbitrary positive irrationals satisfying  $a + b = 1$ .

It is possible to phrase the third theorem in different ways. For instance, we have shown that there is no entire function  $F$  which satisfies

- (i)  $F(z) \leq e^{b|z|}$ ;
- (ii)  $F(2\pi n/b) = -F(2\pi n/a)$ ,  $n = 0, \pm 1, \pm 2, \dots$ ;
- (iii)  $F \in H^2(\mathbf{R})$ .

It has been pointed out that (i) and (ii) are satisfied by the function

$$z \rightarrow \sin(az/2) \sin(bz/2).$$

However, this function is not the Fourier transform of a function in  $L^2(0, b)$ . Indeed, only the function identically equal to zero can satisfy (i) and (iii) while vanishing at all the points  $2\pi n/b$ . Another interpretation arises if we express a function  $f$  in  $L^2(0, b)$  as

$$f(x) = \sum A_n \exp\left(\frac{nx}{b}\right).$$

Then (9) becomes

$$bA_m = - \int_0^b \exp\left(-\frac{mx}{a}\right) \sum A_n \exp\left(\frac{nx}{b}\right) dx$$

or

$$a \left( \exp\left(-\frac{m}{a}\right) - 1 \right) \sum \frac{A_n}{na - mb} = 2\pi i A_m,$$

$m = \pm 1, \pm 2, \dots$  This eigenvalue problem for the doubly infinite matrix has no nontrivial  $l^2$  solutions.

The density of the span of  $\{\Phi_n\}$  gives some insight into the nature of solutions of (E) under the special restrictions placed on  $a$  and  $b$ . If  $F \in L^1(0, 1)$  satisfies (E), then  $F$  must be in the span of linear combinations of the  $\Phi_n$ . There must exist functions  $P_n$  of period 1 such that  $P_n(x/a) + P_n(x/b)$  converges to  $F$  in  $L^1(0, 1)$ . Indeed, a quick calculation shows that for any periodic  $G \in L^1(0, 1)$  the function

$$F(x) = G(x/a) + G(x/b)$$

will be a solution of (E) and

$$f(x) = bG(x/a) + aG(x/b)$$

a solution of (e). It is natural to ask whether all solutions can be obtained using these representations. The answer, unfortunately, is no. It is true that such a representation is unique. For if  $G \in L^1(0, 1)$  is periodic and satisfies

$$G(x/a) + G(x/b) = 0, \quad 0 \leq x \leq 1,$$

we can rewrite this as

$$G(x) + G(\lambda x) = 0, \quad 0 \leq x \leq 1/a.$$

In particular, for  $0 \leq x \leq \lambda^{-1}$  we have

$$G(\lambda x) = G(\lambda x + \lambda),$$

using both properties of  $G$ . Equivalently,

$$G(t) = G(t + \lambda) \pmod{1}$$

and, as  $\lambda$  is irrational,  $G$  is constant. The constant must be 0. The same argument applies to the second representation as well. Whereas the mapping

$$G(x) \rightarrow G(x/a) + G(x/b)$$

has a dense range among the solutions of (E), it is not surjective. For instance, the enigmatic function  $\Phi_0$  cannot be expressed in this manner. Verification of this statement is left to the reader.

4. The amount of information which can be derived from the conclusions of the preceding section must be left for another time. So too must a characterization of extreme solutions of (e). We conclude with two observations regarding the case where  $a$  and  $b$  are irrational.

PROPOSITION 2. *If  $T \in K$  is given by (1) where  $0 < g(x) < 1$  almost everywhere, then  $T\chi_E = \chi_F$  is only possible if  $|E| = 1$  or 0.*

PROOF. This is the same as asking that

$$g(x)\chi_E(\phi(x)) + (1 - g(x))\chi_E(\psi(x)) = \chi_F(x).$$

The condition on  $g$  forces  $\phi^{-1}(E) = \psi^{-1}(E) = F$ . If  $0 \leq x < a$  then  $x \in F$  implies  $x + b \in F$  while  $0 \leq x < b$  with  $x \in F$  implies  $x - a \in F$ . Therefore  $F = (F + b) \pmod{1}$ , completing the argument.

PROPOSITION 3. *A nonconstant characteristic function cannot satisfy (e).*

PROOF. If  $\chi_G$  were such a function we should have

$$a(\chi_G(ax) - \chi_G(ax + b)) = b(\chi_G(bx) - \chi_G(bx + a)),$$

forcing  $\chi_G(ax) = \chi_G(ax + b)$  and  $\chi_G(bx) = \chi_G(bx + a)$  for  $x \in [0, 1)$ . But this leads to the same conclusion as in the proof of the preceding proposition.

Whatever their nature, the extreme solutions of (1) must lie somewhere in between the types described in these assertions.

#### BIBLIOGRAPHY

1. J. Dhombres, *Some aspects of functional equations*, Chulalongkorn University, Bangkok, Thailand, 1979.
2. J. Lindenstrauss, *A remark on extreme doubly stochastic measures*, Amer. Math. Monthly **72** (1965), 379–382.
3. A. Renyi, *Representations for real numbers and their ergodic properties*, Acta Math. Acad. Sci. Hungar. **8** (1957), 477–493.
4. J. V. Ryff, *On the representation of doubly stochastic operators*, Pacific J. Math. **13** (1963), 1379–1386.
5. ———, *Orbits of  $L^1$ -functions under doubly stochastic transformations*, Trans. Amer. Math. Soc. **117** (1965), 92–100.
6. ———, *The functional equation  $af(ax) + bf(bx + a) = bf(bx) + af(ax + b)$ , extensions and almost periodic solutions*, Bull. Amer. Math. Soc. **82** (1976), 325–327.
7. ———, *The functional equation  $F(ax) + F(bx + a) = F(bx) + F(ax + b)$ , entire and almost periodic solutions*, J. Reine Angew. Math. **302** (1978), 116–135.
8. Y. Sakai and T. Shimogaki, *Equimeasurability of functions and doubly stochastic operators*, Kodai Math. Sem. Rep. **24** (1972), 203–211.
9. R. C. Shiflett, *Extreme Markoff operators and the orbits of Ryff*, Pacific J. Math. **40** (1972), 201–206.
10. P. Walters, *Invariant measures and equilibrium states for some mappings which expand distances*, Trans. Amer. Math. Soc. **236** (1978), 121–153.

DIVISION OF MATHEMATICAL SCIENCES, NATIONAL SCIENCE FOUNDATION, WASHINGTON, D.C. 20550