ON THE MINIMAL PROPERTY OF
THE FOURIER PROJECTION

BY
E. W. CHENEY(1), C. R. HOBBY(1)
P. D. MORRIS(2), F. SCHURER(3) AND D. E. WULBERT(2)

Denote by $C$ the Banach space of all continuous, $2\pi$-periodic, real-valued functions defined on $[-\pi, \pi]$, the norm being given by the equation $\|f\| = \max |f(x)|$. Denote by $P_n$ the subspace of $C$ consisting of all trigonometric polynomials of degree $\leq n$. Any bounded linear operator $A: C \to P_n$ such that $Ap = p$ for all $p \in P_n$ is a projection of $C$ onto $P_n$. The Fourier projection is such an operator; it is defined by an equation $F_nf = \sum \langle f, p_k \rangle p_k$, where $p_0, \ldots, p_{2n}$ is any basis for $P_n$ which is orthonormal with respect to the inner-product $\langle f, g \rangle = \int_{-\pi}^{\pi} f(t)g(t) dt$.

It was proved by Lozinski [6] that the Fourier projection is a minimal projection of $C$ onto $P_n$. In other words, the inequality $\|F_n\| \leq \|A\|$ is valid for every projection $A$ of $C$ onto $P_n$. This property of $F_n$ has been discussed by other authors [3, p. 254], [4, p. 154], and [1], and it has been an open question whether any other minimal projection from $C$ onto $P_n$ exists. It is proved below that $F_n$ is the only such projection.

The following conventions are used throughout the paper. An integral written simply $\int f$ or $\int f(t) dt$ stands for $(1/2\pi) \int_{-\pi}^{\pi} f(t) dt$. For each point $x \in [-\pi, \pi]$, an evaluation functional $x$ is defined by the equation $xf = f(x)$, $f$ being an arbitrary member of $C$. The shift operator $T_s$ is defined by $xT_sf = f(x + s)$.

Several elementary results from Fourier analysis are required. If $A$ is any projection of $C$ onto $P_n$ then $F_n = \int T_{-s}AT_s ds$. This means that for each $x$ and for each $f$, $xT_sf = \int xT_{-s}AT_s f ds$. This equation of Berman [1] yields at once the inequality $\|F_n\| \leq \|A\|$. The Fourier projection can be expressed as an integral operator as follows: $xF_n f = \int D_nT_s f$, where $D_n$ (the Dirichlet kernel) is given by:

$$D_n(t) = 1 + 2 \sum_{k=1}^{n} \cos kt = \sin (n + \frac{1}{2})t / \sin \frac{1}{2}t.$$

**Theorem.** If $A$ is a projection of $C$ onto $P_n$ such that $\|A\| \leq \|F_n\|$ then $A = F_n$.

**Proof.** The proof of this theorem depends on a number of lemmas. We first show that the theorem follows from Lemmas 1, 2, and 3. We next prove Lemmas 1

---

(1) Supported by the Air Force Office of Scientific Research.
(2) Supported by the National Science Foundation.
(3) Supported by a NATO Science Fellowship, granted by the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

249
and 2 and show that Lemma 3 will follow from Lemma 4. This portion of the paper uses the techniques of analysis. Lemma 4 is equivalent to the statement that a certain specific polynomial of degree 2n does not have a root which is a primitive \((2n+1)\)th root of unity. We next show that Lemma 4 follows from Lemmas 5 and 6. Lemma 5 essentially states that if \(g(x)\) is a polynomial of degree 2n which has a primitive \((2n+1)\)th root of unity as a root, then a certain combination of the coefficients of \(g\) must be zero. In our case the combination of coefficients is too complicated to calculate directly. It is sufficient, however, to show that the combination of coefficients is nonzero modulo an appropriate prime \(q\), and this is the path we follow. Finally, Lemma 6 assures us of the existence of an appropriate prime.

Suppose that the theorem is false. That is, suppose that there exists a minimal projection \(A\) different from \(F_n\). Write Berman’s equation in the form

\[
F_n = \frac{\alpha + \pi}{2\pi} \left[ \frac{1}{\alpha + \pi} \int_{-\pi}^{\pi} T_{-s} \hat{A}T_s \, ds \right] + \frac{\pi - \alpha}{2\pi} \left[ \frac{1}{\pi - \alpha} \int_{0}^{\pi} T_{-s} \hat{A}T_s \, ds \right].
\]

This expresses \(F_n\) as a convex linear combination of two minimal projections: for brevity we write \(F_n = \theta A_1 + (1 - \theta) A_2\), where \(0 < \theta < 1\). The parameter \(\alpha\) is selected so that \(A_1 \neq F_n\). For the second adjoints of the operators, it is clear that \(F_n^{**} = \theta A_1^{**} + (1 - \theta) A_2^{**}\). Henceforth, we write \(F_n^{**}\) and \(A_i^{**}\) simply as \(F_n\) and \(A_i\). This device serves to extend the operators to a domain which includes all bounded measurable functions \((\dagger)\). By Lemma 1, \(\|F_n\| \leq \|A_1\| + (1 - \theta) \|A_2\| \leq \|F_n\|\).

This shows that \(\hat{A}T_{-\alpha} = \|F_n\|\), where \(A\) denotes either \(A_1\) or \(A_2\). Here \(\sigma\) denotes \(\text{sgn} \, D_n\), the sign of \(D_n\).

**Assertion**. Each functional \(\hat{A}\) is absolutely continuous with respect to Lebesgue measure. For simplicity let \(x = 0\). Then \(\hat{\|A\|} = \|F_n\|\). Let \(J\) be an interval on which \(\sigma\) is \(+1\). We claim that \(\hat{\|A\|} \geq 0\) if \(f \leq 0\) on \(J\) and \(f = 0\) off \(J\). Indeed, if this is false, then there exists an \(f \in C\) for which \(\|f\| = 1\), \(f \leq 0\) on \(J\), \(f = 0\) off \(J\), and \(\hat{\|A\|} < 0\). Then \(\|\sigma - f\| = 1\) and a contradiction results: \(\|A\| \geq \|\hat{\|A\|}\| \geq \hat{\|A\|} = \|F_n\| \geq \|\hat{\|A\|} - \hat{\|A\|} F_n\|\).

In the same way, if \(J\) is an interval on which \(\sigma\) is \(-1\) then \(\hat{\|A\|} \leq 0\) for every \(f\) such that \(f \geq 0\) on \(J\) and \(f = 0\) off \(J\). Now let \(S\) be any set of (Lebesgue) measure zero, and let \(h\) be its characteristic function. It is to be proved that \(\hat{\|A\|} h = 0\). It is sufficient to establish this under the additional assumption that \(S\) is contained in an interval \(J\) on which \(\sigma\) is constant, because the entire interval \([-\pi, \pi]\) is the union of a finite number of such subintervals. Suppose for definiteness that \(\sigma\) is

\(\dagger\) Throughout this paper, “measurable sets” and “measurable functions” are understood to be Borel sets and Borel functions. Also Lebesgue measure is understood to be restricted to the Borel sets.
+1 on $J$. Then $\dot{\theta}Ah \geq 0$ by the previous analysis. Hence $0 = \int D_{\theta}Ah = \dot{\theta}F_{\theta}h = \theta \dot{\theta}A_{\theta}h + (1-\theta)\dot{\theta}A_{\theta}h \geq 0$, and $\dot{\theta}Ah = 0$.

Since the range of $A$ is $P_n$, $A$ has a representation in the form $Af = \sum \phi_i(f)p_i$, where $\{p_0, \ldots, p_{2n}\}$ is a basis for $P_n$, and $\phi_i \in C^*$. If $h$ is the characteristic function of a set of measure zero (and if $\phi_i$ is extended to operate on all bounded measurable functions), then $0 = \dot{\theta}Ah = \sum \phi_i(h)p_i(x)$. By linear independence, $\phi_i(h) = 0$ for all $i$.

This proves that each functional $\phi_i$ is absolutely continuous with respect to Lebesgue measure. By the Radon-Nikodym Theorem there exist functions $h_i \in L^1[-\pi, \pi]$ such that $\phi_i(f) = \int fh_i$ for all $f \in C$. Thus $A$ is of the form

$$\dot{\theta}Af = \int K(x, t)f(t) \, dt$$

where $K(x, t) = \sum h_i(x)p_i(x)$.

From previous equations and Lemma 1 we have

$$\|\dot{\theta}A\| = \int |K(x, t)| \, dt = \dot{\theta}AT_{\theta}x = \int K(x, t)\sigma(t-x) \, dt.$$ 

Since $\sigma$ is almost everywhere of magnitude 1, we conclude that for all $x$ and for almost all $t$, $K(x, t)\sigma(t-x) \geq 0$. Since these arguments are valid for both $A_1$ and $A_2$, we obtain integral representations (and kernels $K_1, K_2$) for both. It follows that $D_{\theta}(t-x) = \theta K_1(x, t) + (1-\theta)K_2(x, t)$. Since $K_i(x, t)\sigma(t-x) \geq 0$ for almost all $t$ and $|D_{\theta}| \leq 2n+1$, we see that $|K_i(x, t)| \leq (2n+1)/|\theta(1-\theta)|$. Thus for fixed $x$, $K_i(x, t)$ is a bounded $L^1$-function, and hence an element of $L^2[-\pi, \pi]$.

Now let $v$ be any root of $D_n$, and put $g(x, t) = K(x-v, t)$. Then $g$ (or possibly $-g$) satisfies the hypotheses of Lemma 2. From that lemma, we obtain $g(t, t) = 0$ for almost all $t$, or in terms of $K$, $K(t-v, t) = 0$ for almost all $t$. Thus, for almost all $t$, the trigonometric polynomial $K(x, t)$ has roots (in $x$) at the roots of $D_n(t-x)$.

Since these $2n$ roots determine an element of $P_n$ exactly except for a factor, we must have $K(x, t) = \beta(t)D_n(t-x)$ for an appropriate $\beta \in L^2[-\pi, \pi]$. Clearly $\beta \geq 0$. Let $\alpha = \beta - 1$.

If $p \in P_n$ then $Ap = p$ and consequently $p(x) = \dot{\theta}Ap = \int p \cdot (1+\alpha) \cdot T_{\theta}x \, D_n = \dot{\theta}F_n(p + ap)$. This proves that $F_n(ap) = 0$ for all $p \in P_n$. If $q \in P_n$ then $0 = \langle q, ap \rangle = \langle qp, a \rangle$. Since $qp$ ranges over all of $P_{2n}$, we conclude that $\alpha \perp P_{2n}$.

Next we observe that $\alpha \perp T_x|D_n|$ for all $x$. This follows from the equation $\|F_n\| = \|\dot{\theta}A\| = \int |K(x, t)| \, dt = \int |(1+\alpha) \cdot T_{\theta}x \, D_n| = \|F_n\| + \int \alpha \cdot T_{\theta}x \, |D_n|$.

Since $\alpha \in L^2[-\pi, \pi]$ we have $\alpha = \sum_{k=-\infty}^{\infty} c_k e_k$, where $e_k(x) = e^{ikx}$. Also let $|D_n| = \sum d_k e_k$. Then $T_x|D_n| = \sum d_k e_k(x)e_k$. Since $\alpha$ is orthogonal to this function, we have $\sum c_k d_k e_k(x) = 0$. Hence $c_k d_k = 0$ for all $k$. But $\alpha \perp P_{2n}$, and hence $c_k = 0$ for $|k| \leq 2n$. By Lemma 3, $d_k \neq 0$ when $|k| > n$, and hence $c_k = 0$ for $|k| > n$. Hence $c_k = 0$ for all $k$. It follows that $\alpha = 0$, $\beta = 1$, $K(x, t) = D(t-x)$, and $A = F_n$. Thus the proof will be complete as soon as we verify Lemmas 1, 2, and 3.
Lemma 1. Let $T$ be a bounded linear operator from $C$ into $P_n$, and let $x$ be a point. If there exists an element $g \in L^1$ such that $\hat{x}Tf = \int g f$ for all $f \in C$ then $\hat{x}T^*f = \int g f$ for all bounded measurable $f$.

Proof. Let $L_0^\infty$ denote the normed linear space of bounded measurable functions with supremum norm. Let $\phi(f) = \int g f$, and let $w$ denote the $w(L_0^\infty, C^*)$-topology on $L_0^\infty$. We shall prove that $\phi$ and $\hat{x}T^*$ are $w$-continuous extensions of $xT$, and that $C$ is $w$-dense in $L_0^\infty$. If $f_\epsilon$ is a net in $L_0^\infty$ such that $f_\epsilon \to 0$ (mod $w$) then $\langle g, f_\epsilon \rangle \to 0$ because $\langle g, \cdot \rangle \in C^*$. Thus $\phi$ is $w$-continuous. Since the adjoint of any bounded operator is continuous in the two weak* topologies, [2, p. 478], we see that $T^*$ is $w(C^*, C^*) - w(C^*, C^*)$-continuous. Since $f_\epsilon \in C^*$, $T^*f_\epsilon \to 0$ in the $w(C^*, C^*)$-topology. Since $\hat{x} \in C^*$, $\hat{x}T^*f_\epsilon \to 0$. Thus $\hat{x}T^*$ is $w$-continuous. By Goldstine's Theorem [2, p. 424], $C$ is $w(C^*, C^*)$-dense in $C^*$. Hence it is also $w$-dense in $L_0^\infty$.

Lemma 2. Let $g$ be a bounded function of two variables such that for each $x$, $g(x, \cdot) \in L^1[-\pi, \pi]$, and for each $t$, $g(\cdot, t) \in P_n$. Let $\delta > 0$, and assume that for each $x$ and for almost all $t \in (0, \delta)$, $g(x, x-t) \leq g(x, x+t)$. Then $g(x, x) = 0$ for almost all $x$.

Proof. Assume that the conclusion is false. Without loss of generality, let $g(x, x) > 0$ on a set of positive measure. Then for some $\epsilon > 0$, $g(x, x) \geq 2\epsilon$ on a set of positive measure. By the regularity of Lebesgue measure, there is a closed set $S$ of positive measure such that $g(x, x) \geq 2\epsilon$ for all $x \in S$. Since $g(\cdot, t)$ forms a bounded set in $P_n$, there exists a common modulus of continuity for all these functions. Hence there is a number $\theta > 0$ such that $g(z, x) \geq \epsilon$ whenever $x \in S$ and $x \leq z \leq x + \theta$. Let $T = \{(z, x) : x \in S$ and $x \leq z \leq x + \theta\}$. Let $f$ denote the characteristic function of $T$. By the Fubini Theorem, $\int g f = \int \int g(x, t)f(x, t) \, dt \, dx$, the integral on the left being over the square $[-\pi, \pi]^2$. Since $g \geq \epsilon$ on $T$, and $T$ has measure $\theta \mu(S) > 0$, we have $\int g f \geq 0$. Hence there is an $x$ for which $\int g(x, t)f(x, t) \, dt > 0$. Since $f(x, t) = 0$ if $t > x$ or $t < x - \delta$, we have $\int_x^{x+\delta} g(x, t)f(x, t) \, dt > 0$. But this is not possible since by hypothesis $g(x, t) \leq 0$, for almost all $t$ in $(x-\delta, x)$.

Lemma 3. If the Fourier series of $|D_n|$ is $\sum_{n=-\infty}^n d_k e_k$ then $d_k \neq 0$ when $|k| > n$.

Proof. Since $|D_n|$ is periodic and even, we have

$$d_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_n(t)| e^{-ikt} \, dt = \frac{1}{\pi} \int_0^{2\pi} \left| \sum_{n=0}^n \cos jt \right| \cos kt \, dt.$$ 

Here the prime indicates that the term corresponding to $j=0$ is to be halved. Since $D_n(t) = \sin (n+\frac{1}{2})t / \sin \frac{1}{2}t$, it changes sign at each of the points $t_1, \ldots, t_{2n}$, where $t_v = \pi n / (n+\frac{1}{2})$. Hence

$$d_k = \frac{1}{\pi} \sum_{v=0}^{2n} (-1)^v \int_{t_v}^{t_v+\frac{1}{2}} \left| \sum_{j=0}^n \cos jt \right| \cos kt \, dt$$

$$= \frac{1}{2\pi} \sum_{v=0}^{2n} (-1)^v \int_{t_v}^{t_v+\frac{1}{2}} \left[ \cos (k+j)t + \cos (k-j)t \right] \, dt.$$
Now assume that \( k > n \) so that \( k+j \neq 0 \) and \( k-j \neq 0 \). Then

\[
d_k = \frac{1}{2\pi} \sum_{j=0}^{n} \left( \sum_{v=0}^{2n} (-1)^v \left[ \frac{\sin (k+j)t}{k+j} + \frac{\sin (k-j)t}{k-j} \right] \right) t_{v+1}
\]

\[
= \frac{1}{2\pi} \text{Im} \sum_{j=0}^{n} \left( \sum_{v=0}^{2n} (-1)^v \left[ \frac{\beta^{(k+j)(v+1)} - \beta^{(k+j)v}}{k+j} + \frac{\beta^{(k-j)(v+1)} - \beta^{(k-j)v}}{k-j} \right] \right)
\]

where \( \beta = e^{i\pi (n+1/2)} \). After carrying out the \( v \)-summation and replacing \( \beta^{2n+1} \) by 1 we obtain

\[
d_k = \frac{1}{\pi} \text{Im} \sum_{j=0}^{n} \left( \sum_{v=0}^{2n} (-1)^v \left[ \frac{\beta^{k+j} - 1}{(k+j)(\beta^{k+j} + 1)} + \frac{\beta^{k-j} - 1}{(k-j)(\beta^{k-j} + 1)} \right] \right)
\]

\[
= \frac{1}{\pi} \text{Im} \sum_{j=0}^{n} \frac{\beta^j - 1}{(k+j)(\beta^j + 1)}
\]

\[
= \frac{1}{\pi} \text{Im} \sum_{j=0}^{k+n} \frac{1}{j} \left( 1 - \frac{2}{\beta^j + 1} \right)
\]

\[
= -\frac{2}{\pi} \text{Im} \sum_{j=0}^{k+n} \frac{1}{j(\beta^j + 1)} = -\frac{1}{\pi} \sum_{j=0}^{k+n} \frac{1}{j} \left( \beta^j + 1 - \beta^{-j} + 1 \right)
\]

\[
= \frac{1}{\pi} \sum_{j=0}^{k+n} \frac{\beta^j - 1}{\beta^j + 1}
\]

The proof is completed by an appeal to Lemma 4.

**Lemma 4.** If \( k \) and \( n \) are positive integers then

\[
\sum_{v=k}^{k+2n} \frac{1}{v} \beta^v - 1 \neq 0.
\]

**Proof.** We first note that the term in (1), involving the unique \( v \) such that \( 2n+1 \) divides \( v \), is zero. Hence this term may be deleted from the summation. Let \( d(v) = (v, 2n+1) \), the greatest common divisor of \( v \) and \( 2n+1 \). Since \( \beta \) is a primitive \((2n+1)\)th root of unity, \( \beta^v \) is a primitive \(((2n+1)/d(v))\)th root of unity. We set \( s(v) = (2n+1)/d(v) - 1 \) and note that \( \beta^v - 1 = -\beta^v + 1 \sum_{t=1}^{s(v)} \beta^{-it} \).

Thus if (1) fails to hold we must have

\[
\sum_{v=k}^{k+2n} \frac{1}{v} \beta^v - 1 \neq 0.
\]

Using only the relation \( \beta^{2n+1} = 1 \) we rewrite (2) as

\[
\sum_{i=0}^{2n} d_i \beta^i = 0.
\]
where for each $i$

$$d_i = \sum_{v = k; 2n + 1 | v}^{k + 2n} \delta_i(v)$$

and

$$\delta_i(v) = \begin{cases} 0, & \text{if there is no } t \in \{1, \ldots, s(v)\} \text{ satisfying } vt \equiv i(2n + 1) \\ -1, & \text{if the solution } t \in \{1, \ldots, s(v)\} \text{ to } vt \equiv i(2n + 1) \text{ is odd} \\ 1, & \text{if the solution is even.} \end{cases}$$

It is a consequence of Lemma 5 (below) that a certain combination of the $d_i$'s is equal to zero. In order to apply Lemma 5 we first factor $2n + 1$. Suppose $2n + 1 = p_1^{a_1}p_2^{a_2} \cdots p_r^{a_r}$ where the $p_i$ are distinct primes. Let $m_i = \prod_{j \neq i} p_j^{a_j}$. The greatest common divisor of $\{m_j : j = 1, \ldots, r\}$ is 1, therefore there exist integers $c_j$ such that $\sum c_j m_j = 1$. Now let $a_j = \beta^{c_j} m_j$. Clearly $\beta = a_1 a_2 \cdots a_r$. Moreover, since $\beta$ is a primitive $(2n + 1)$th root of unity, each $a_j$ is a primitive $p_j^{a_j}$th root of unity. For each $i$, let $i_j$ denote the element of $0, 1, \ldots, p_j^{a_j} - 1$ which is congruent $i$ modulo $p_j^{a_j}$. Then $\beta^i = a_1^{i_1} a_2^{i_2} \cdots a_r^{i_r}$, and we can write (3) as

$$\sum_{0 \leq i_1, \ldots, i_r, i = 1, \ldots, r} b_{(i_1, \ldots, i_r)} a_1^{i_1} \cdots a_r^{i_r} = 0,$$

where $b_{(i_1, \ldots, i_r)}$ denotes the (unique) $d_i$ for which $\beta^i = a_1^{i_1} \cdots a_r^{i_r}$. We see from Lemma 5 that (4) implies

$$\sum_{s = 0}^{r} (-1)^s D_s = 0,$$

where $D_s$ is the sum of those coefficients $b$, whose vector subscripts (i) have exactly $s$ nonzero entries and (ii) are such that if $i_j \neq 0$ then $i_j = p_j^{a_j - 1}$. We will complete the proof of Lemma 4 by showing that (5) is a contradiction.

We see from (3) that each $D_s$ is a sum of terms $\delta_i(v)/v$. Suppose that we can find a prime $q$ such that if $q^m$ is the highest power of $q$ which divides an element of $S = \{k, k + 1, \ldots, k + 2n\}$ then,

(i) there is a unique $\mu \in S$ which is divisible by $q^m$, say $\mu = q^m \mu_0$, and

(ii) this $\mu$ is not divisible by $2n + 1$.

Then, for each $v \in S$, the reduced form of $q^m/v$ will have a denominator prime to $q$. Moreover, $q^m/v \equiv 0$ (modulo $q$) except when $v = \mu$, and $q^m/\mu = 1/\mu_0 \neq 0$ ($q$). If (5) holds, then, in particular,

$$\sum_{s = 0}^{r} (-1)^s q^m D_s \equiv 0 \pmod{q}.$$

The only terms in the left-hand side of (7) which are nonzero modulo $q$ are the terms $\delta_i(\mu)/\mu_0$. We shall count the number of summands $\delta_i(\mu)/\mu$ in $D_s$ for which
\( \delta_i(\mu) \) is nonzero. We are only concerned with \( s \neq 0 \) since it is clear from (3) that \( D_0 = 0 \). Suppose that the greatest common divisor of \( \mu \) and \( 2n+1 \) is \( p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r} \). Clearly, \( 0 \leq m_i \leq n_i \). Since \( 2n+1 \) does not divide \( \mu \) we know that \( m_i < n_i \) for some \( i \in \{1, \ldots, r\} \). There is no loss of generality if we suppose that \( m_i < n_i \) for \( i = 1, \ldots, h \), and \( m_j = n_j \) for \( j > h \). We claim that the number of summands \( \delta_i(\mu) / \mu \) in \( D_0 \) for which \( \delta_i(\mu) \neq 0 \) is exactly \( C_{h,s} \). [We set \( C_{h,s} = 0 \) when \( s > h \) otherwise \( C_{h,s} = \binom{h}{s} \).]

By the definition of \( D_0 \), we are only concerned with those \( i \) such that \( \beta^i = a_1^{i_1} \cdots a_r^{i_r} \) where exactly \( s \) of the \( i_j \) are nonzero and each \( i_j \) is either 0 or \( p_j^{n_j-1} \). Furthermore, \( \delta_i(\mu) \) is nonzero if and only if there is a solution \( x \) to the congruence \( \mu x \equiv i (2n+1) \). Hence \( \delta_i(\mu) \) is nonzero if and only if the greatest common divisor of \( \mu \) and \( 2n+1 \) divides \( i \). If we have \( i_j = p_j^{n_j-1} \) for some \( j > h \) then clearly \( p_j^{n_j} \) does not divide \( i \), hence \((\mu, 2n+1) \mid i \) and it follows that \( \delta_i(\mu) = 0 \). On the other hand, if \( i_j = 0 \) for every \( j > h \) and if \( i_j \) is either \( p_j^{n_j-1} \) or 0 for \( j \leq h \), then \( i \) is divisible by \((\mu, 2n+1) \) and it follows that \( \delta_i(\mu) \neq 0 \). We have shown that \( \delta_i(\mu) / \mu \) makes a nonzero contribution to \( D_0 \) if and only if \( \beta^i = a_1^{i_1} \cdots a_r^{i_r} \) where exactly \( s \) of the \( i_j \) are \( p_j^{n_j-1} \), \( j \in \{1, \ldots, h\} \), and all other \( i_j \) are zero. There are clearly exactly \( C_{h,s} \) such values of \( i \).

We have shown that the left-hand side of (7) consists, modulo \( q \), of \( N = C_{h,1} + C_{h,2} + \cdots + C_{h,h} \) terms each of which is congruent \( \pm 1 / \mu_0 \). If (7) holds then there must be a sum of \( N \) terms \( \pm 1 \) which is either 0 or a nonzero multiple of \( q \). Since \( N = 2^h - 1 \) is odd, the sum cannot be zero or a multiple of 2. If \( q \) is greater than \( N \) then clearly the sum cannot be a nonzero multiple of \( q \). That is, (6) is a contradiction if \( q = 2 \) or if \( q > 2^h \). Since \( h \leq r \) and \( 2n+1 \) involves \( r \) distinct primes we see that the proof is complete if \( q = 2 \) or if \( q > 2n+1 \). We now appeal to Lemma 6. If \( \mu \) is the (unique) element of \( S \) which is divisible by a maximal power of 2, and if \( 2n+1 \) does not divide \( \mu \), then we take \( q = 2 \) and the proof is complete. On the other hand, if \( 2n+1 \) divides \( \mu \) we let \( q \) be the prime \( p \) of Lemma 6. Then \( q > 2n+1 \) and \( q \) divides \( v \in S \) where \( v \) is not a multiple of \( 2n+1 \). The pair \( q, v \) clearly satisfies (6). Thus the proof is again complete.

**Lemma 5.** Suppose that \( p_1, p_2, \ldots, p_r \) are distinct primes, and \( n_1, n_2, \ldots, n_r \) are positive integers, and that \( a_i \) is a primitive \( p_i^{n_i} \)th root of unity for \( i = 1, 2, \ldots, r \). Let \( b_{i_1, \ldots, i_r} \) be rational numbers such that

\[
\sum b_{i_1, \ldots, i_r} a_1^{i_1} \cdots a_r^{i_r} = 0,
\]

where we sum over \( 0 \leq i_1 < p_1^{n_1}, \ldots, 0 \leq i_r < p_r^{n_r} \). Then

\[
\sum_{s=0}^{r} (-1)^s D_s = 0,
\]

where \( D_s \) is the sum of those coefficients \( b \) whose vector subscripts \( (i_1, \ldots, i_r) \) have exactly \( s \) nonzero entries and are such that if \( i_j \neq 0 \), then \( i_j = p_j^{n_j-1} \).

**Proof.** The proof is by induction on \( r \). The following argument provides the
induction step and at the same time serves as the proof for \( r=1 \). For each \( j=0, \ldots, p_1^{k_1} - 1 \) we set

\[ C_j = \sum b_{i_0,i_2,\ldots,i_r} a_{i_0}^{i_2} \cdots a_{i_r}^{i_r}, \]

where \( 0 \leq i_2 < p_2^{k_2}, \ldots, 0 \leq i_r < p_r^{k_r} \). Let \( g(a_1) = 0 \) and hence \( g(x) \) is divisible by the minimal polynomial \( t(x) \) for \( a_1 \) over \( F = \mathbb{F}(a_2, \ldots, a_r) \), where \( \mathbb{F} \) denotes the rational numbers. Since \( a_1 \) is a primitive \( p_1^{k_1} \)th root of unity, its minimal polynomial over \( F \) is the same as its minimal polynomial over \( \mathbb{F} \), [Kronecker, Werke, Vol. 1, Leipzig, 1895, p. 85] namely,

\[ t(x) = 1 + x^{p_1^{k_1} - 1} + x^{2p_1^{k_1} - 1} + \cdots + x^{(p_1 - 1)p_1^{k_1} - 1}. \]

We know that \( g(x) = t(x)h(x) \), where \( h(x) \) is a polynomial over \( F \) of degree \( \deg(g(x)) - \deg(t(x)) \).

Hence

\[ \deg(h(x)) \leq (p_1^{k_1} - 1) - (p_1 - 1)p_1^{k_1} - 1 = p_1^{k_1} - 1 - 1. \]

Letting \( h_0 \) denote the constant term of \( h(x) \), we see that the coefficients of \( 1, x^{p_1^{k_1} - 1}, x^{2p_1^{k_1} - 1}, \ldots, x^{(p_1 - 1)p_1^{k_1} - 1} \) in \( t(x)h(x) \) are all equal to \( h_0 \). Therefore,

\[ C_0 = C_{p_1^{k_1} - 1} = \cdots = C_{(p_1 - 1)p_1^{k_1} - 1} = h_0. \]

In particular, \( C_0 - C_{p_1^{k_1} - 1} = 0 \), that is

\[ \sum (b_{0,i_2,\ldots,i_r} - b_{i_0,i_2,\ldots,i_r}) a_{i_0}^{i_2} \cdots a_{i_r}^{i_r} = 0, \]

where \( 0 \leq i_2 < p_2^{k_2}, \ldots, 0 \leq i_r < p_r^{k_r} \). In particular, if \( r=1 \), then \( b_{00} = b_{i_0,i_2,\ldots,i_r} \), which proves the lemma for \( r=1 \).

**Remark.** One can easily modify the proof of Lemma 5 to obtain additional relations on the coefficients of a polynomial of degree \( \leq m - 1 \) which has an \( m \)th root of unity as a root. We actually proved that \( C_i = C_{i + p_1^{k_1} - 1} = C_{i + 2p_1^{k_1} - 1} = \cdots \), but we only used the case for \( i=0 \). The natural generalization of Lemma 5 yields a system of linear equations which can be used to determine the coefficients of the minimal polynomial for an \( m \)th root of unity.

**Lemma 6.** Let \( S = \{k, k+1, \ldots, k+2n\} \) where \( n \) and \( k \) are positive integers. Let \( \mu \) be the (unique) element of \( S \) which is divisible by a maximal power of 2. Then

(i) \( 2n+1 \) does not divide \( \mu \), or

(ii) \( 2n+1 \) divides \( \mu \) and there is a prime \( p > 2n+1 \) and an element \( v \neq \mu \) of \( S \) such that \( p \) divides \( v \).

**Proof** (5). First, suppose that \( 2n+1 \) is an element of \( S \). Then \( 2n+1 \) cannot

(5) This proof was suggested by a paper of Erdős [5]. We are indebted to John Selfridge for calling our attention to [5].
divide \( \mu \) unless \( 2n + 1 = \mu \). Since \( \mu \) is even this completes the proof for \( 2n + 1 \in S \).

Thus we may suppose that \( k > 2n + 1 \).

There is nothing to prove unless \( 2n + 1 \) divides \( \mu \). If \( 2^m \) is the highest power of \( 2 \) dividing \( \mu \), then \( 2^{m+1} > 2n + 1 \), since \( 2^{m+1} \) does not divide any element of \( S \). By the theorem of Sylvester and Schur (see [5]) there is a prime \( p > 2n + 1 \) such that \( p \) divides some element of \( S \). If \( p \) divides \( v \) where \( v \neq \mu \) then there is nothing to prove. Thus we may suppose that \( \mu = (2n + 1)2^m \mu_0 \) where \( \mu_0 \) is an integer. Therefore, \( \mu > (2n+1)^3/2 \), hence \( k + 2n > (2n+1)^3/2 \). We suppose that the lemma is false. Then \( \mu \) is the only element of \( S \) which is divisible by a prime greater than \( 2n + 1 \). For each prime \( p_i \leq 2n + 1 \) we let \( p_i^{q_i} \) denote the highest power of \( p_i \) which divides \( C_{k+2n,2n+1} \). It follows that

\[
C_{k+2n,2n+1} \leq \mu \prod p_i^{q_i}.
\]

As was shown in [5], \( p_i^{q_i} \leq k + 2n \). Let \( d = d(n) \) denote the number of primes which do not exceed \( 2n + 1 \). Then

\[
C_{k+2n,2n+1} \leq \mu (k + 2n)^d \leq (k + 2n)^d + 1.
\]

For each \( t = 1, \ldots, 2n \), we have \( (k + 2n)/(2n + 1) < (k + 2n - t)/(2n + 1 - t) \), therefore

\[
\left(\frac{k + 2n}{2n + 1}\right)^{2n+1} < (k + 2n)^{d+1}.
\]

If \( n \geq 4 \), then \( d(n) \leq n \). Therefore, \( (k + 2n)^n < (2n+1)^{2n+1} \) for \( n \geq 4 \). Using \( (k + 2n) > (2n+1)^3/2 \) we have a contradiction for \( n \geq 4 \).

The remaining cases \( (n = 1, 2, 3) \) can be disposed of by routine arguments. We shall sketch the proof for \( n = 3 \). Recall that we are only concerned with the situation where (i) \( k > 2n + 1 \), and where (ii) \( \mu \) is divisible by \( 2^m \), \( 2n + 1 \), and by every prime \( p > 2n + 1 \) which divides an element of \( S \).

When \( 2n + 1 = 7 \), the six elements of \( S \) have no prime factors other than \( 2, 3, \) and \( 5 \). Let \( v_1 \) and \( v_2 \) denote, respectively, the elements of \( S \) divisible by maximal powers of \( 3 \) and \( 5 \). Let \( T = S \setminus \{\mu, v_1, v_2\} \), and let \( t \) denote the product of elements of \( T \). Then \( T \) contains at least four elements, each of which is greater than \( 7 \) (since \( k > 2n + 1 \)). Therefore \( t > 7^4 \). This is a contradiction since one can show that \( t \leq 2^4 \cdot 3^2 \cdot 5 \). We show, for example, that \( 2^5 \) does not divide \( t \). Since \( \mu \notin T \) and \( \mu \) is divisible by the maximal power of \( 2 \), we know that no element of \( T \) is a multiple of \( 8 \). At most one element of \( T \) can be a multiple of \( 4 \), and at most two additional elements of \( T \) are divisible by \( 2 \). Thus \( t \) is not divisible by \( 2^5 \). Similar arguments show that \( t \) is not divisible by \( 3^3 \) or by \( 5^2 \).

The cases for \( 2n + 1 = 3, 5 \) can be handled by arguments similar to the above.

**Added in Proof.** Since the submission of this manuscript the doctoral dissertation of P. V. Lambert [7] has come to our attention. Dr. Lambert has shown that if \( C \) is taken to be the complex-valued continuous \( 2\pi \)-periodic functions, then any minimal projection onto the \( n \)th order trigonometric polynomials is necessarily
real. Thus the minimal norm projection is also unique in this case. Among many other results Dr. Lambert gives examples of compact Abelian groups for which the natural (Fourier) projections onto subspaces generated by characters are not unique, although minimal.

REFERENCES