C.R. Math. Rep. Acad. Sci. Canada - Vol. IV, No. 4, August 1982 Août

C.R. Math. Rep. Acad. Sci. Canada - Vol. IV, No. 4, August 1982 Août

ON THE ZEROS OF EXPONENTIAL POLYNOMIALS

László Székelyhidi

Presented by J. Aczél, F.R.S.C.

ABSTRACT: It is proved that the set of zeros of a nonzero continous exponential polynomial on a locally compact Abelian group generated by any neighborhood of zero has measure zero.

It is a classical result that the set of roots of a complex polynomial of n variables has measure zero ([3]) . It is also well known ([2]), that the set of zeros of a trigonometric polynomial on a locally compact connected Abelian group has measure zero. In this paper we extend these results to exponential polynomials on some types of locally compact Abelian groups.

If G is an Abelian group and X is a linear space over the complex field $\mathbb{C}$, then we use the difference operators defined on functions $f: G \to X$ by

$$\Delta_{y_1,\dots,y_n} f(x) = \sum_{1 \le i_1 < \dots < i_k \le n} (-1)^{n-k} f(x+y_{i_1}+\dots+y_{i_k})$$

for all $x, y_1, \dots, y_n$ in G. (For $k=0$ the sum is $f(x)$.) In particular we write

$$\Delta_y^n f(x) = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} f(x+ky)$$

for all x, y in G. A function $f: G \to X$ is called a polynomial of degree at most n if

$$\Delta_{y_1,\dots,y_{n+1}} f(x) = 0$$

for all $x, y_1, \dots, y_{n+1}$ in G. It is well known ([1]), that each polynomial f of degree at most n has a representation

$$f(x) = \sum_{k=0}^{n} A_k^{(k)}(x)$$

where $A_k : G^k \to X$ is a k-additive, symmetric function and $A_k^{(k)}$ is the diagonalization of A_k defined by $A_k^{(k)}(x) = A_k(x_1, \dots, x_k)$ with $x_1 = \dots = x_k = x$. /Here $G^0 = G$ and A_0 is a constant./ Further, the relation

$$A_n(y_1, \dots, y_n) = \frac{1}{n!} \Delta_{y_1, \dots, y_n} f(x)$$

is valid for all $x, y_1, \dots, y_n$ in G.

A function $m : G \to \mathbb{C}$ is called an exponential function, if it is a homomorphism of G into the multiplicative group of nonzero complex numbers. A function $f : G \to X$ is called an exponential polynomial if it has a representation

$$f(x) = \sum_{i=1}^{n} m_i(x) p_i(x)$$

where $p_i : G \to X$ is a polynomial and $m_i : G \to \mathbb{C}$ is an exponential function for each i.

<u>THEOREM</u> 1. <u>Let G be a topological Abelian group which is generated by any neighborhood of zero and let X be a complex linear space. If a polynomial $p : G \to X$ vanishes on some nonvoid open set, then it vanishes everywhere.</u>

<u>PROOF</u>. Let $p(x) = \sum_{k=0}^{n} A_k^{(k)}(x)$ for all x in G, where $A_k : G^k \to X$ is k-additive and symmetric. It is enough to show that $A_n = 0$ and we may obviously assume that p vanishes on some neighborhood U of zero. Let $V \subset U$ be a neighborhood of the zero such that $x, y_1, \dots, y_n$ are in V implies $x + y_1 + \dots + y_n$ is in U. Then

$$A_n(y_1, \dots, y_n) = \frac{1}{n!} \Delta_{y_1, \dots, y_n} p(x) = \frac{1}{n!} \sum_{1 \le i_1 < \dots < i_k \le n} (-1)^{n-k} p(x + y_{i_1} + \dots + y_{i_k}) = 0$$

holds for all $x, y_1, \dots, y_n$ in V, that is A_n vanishes on V^n. Using the multi-additivity of A_n and the fact that V^n generates G^n we have $A_n = 0$ and the theorem is proved.

THEOREM 2. Let G be a locally compact Abelian group which is generated by any neighborhood of the zero and let X be a complex linear space. If a polynomial $p: G \to X$ vanishes on a measurable set of positive measure, then it vanishes everywhere.

PROOF. Using the notations of the preceding theorem, there exists a compact set $K \subset G$ with $\lambda K > 0$ for which $p(x) = 0$ for all x in K (λ denotes a fixed Haar-measure on G). By the theorem of Steinhaus the function $y \to \lambda(K \cap K-y \cap \dots \cap K-ny)$ is continuous on G and as it is positive at zero, there exits a neighborhood $U \subset G$ of zero such that $K \cap K-y \cap \dots \cap K-ny$ is nonvoid for all y in U. On the other hand

$$A_n^{(n)}(y) = \frac{1}{n!} \Delta_y^n p(x) = \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} p(x+ky) = 0$$

whenever x belongs to $K \cap K-y \cap \dots \cap K-ny$, since $x+ky$ belongs to K $(k=0,1,\dots,n)$. It means that $A_n^{(n)}$ vanishes on U. Then by the preceding theorem our statement follows.

THEOREM 3. Let G be a topological Abelian group which is generated by any neighborhood of the zero and let X be a complex linear space. If an exponential polynomial $f: G \to X$ vanishes on some nonvoid open set, then it vanishes everywhere.

PROOF. Let $f(x)=\sum_{i=1}^{n} m_i(x)p_i(x)$ for all x in G, where $m_i : G \to \mathbb{C}$ is an exponential function, $p_i : G \to X$ is a polynomial and $m_i \neq m_j$ if $i \neq j$. We prove the theorem by induction on n. For $n=1$ it follows from theorem 1 since $m_i(x) \neq 0$ for all x in G $(i=1,\dots,n)$.

Suppose that it is proved for $n \leq k$ and let $n=k+1$. Let f vanish on a neighborhood $U \subset G$ of zero. Let $p_i(x)=A_i^{(\mu_i)}(x)+q_i(x)$ for x in G, where $A_i : G^{\mu_i} \to X$ is μ_i-additive and symmetric and $q_i : G \to X$ is a polynomial of degree at most $\mu_i - 1$ $(i=1,\dots,k+1)$. It is obviously enough to show that $A_2^{(\mu_2)}=0$. Let

$$\varphi(x)=p_1(x)+\sum_{i=2}^{k+1} m_i(x)m_1(x)^{-1}p_i(x)$$

for all x in G, then φ vanishes on U. Let $V \subset U$ be a neighborhood of zero with the property that $x+ky$ belongs to U whenever x,y is in V $(k=0,1,\dots,N+1)$ where N is the degree of p_1. Then

$$\Delta_y^{N+1}\varphi(x)=\Delta_y^{N+1}p_1(x)+\sum_{i=2}^{k+1} m_i(x)m_1(x)^{-1}\left[\left(m_i(y)m_1(y)^{-1}-1\right)^{N+1}A_i^{(\mu_i)}(x)+q_{i,y}(x)\right]=0$$

for all x,y in V, where $q_{i,y} : G \to X$ is a polynomial of degree at most $\mu_i - 1$ $(i=2,\dots,k+1)$. It follows that

$$\sum_{i=2}^{k+1} m_i(x)\left[\left(m_i(y)m_1(y)^{-1}-1\right)^{N+1}A_i^{(\mu_i)}(x)+q_{i,y}(x)\right]=0$$

whenever x,y is in V. As $m_1 \neq m_2$, there exists an element y in V such that $m_2(y)m_1(y)^{-1} \neq 1$. Indeed, otherwise the exponential property of m_1 and m_2 with the fact that V generates G would imply that $m_1=m_2$.

Substituting this element into the above equation, we have by induction $(m_2(y)m_1(y)^{-1}-1)^{N+1} A_2^{(\mu_2)}(x) + q_{2,y}(x) = 0$ for all x in G. As the coefficient of $A_2^{(\mu_2)}$ differs from zero and the degree of $q_{2,y}$ is at most μ_2-1, hence $A_2^{(\mu_2)}=0$ and the theorem is proved.

THEOREM 4. *Let G be a locally compact Abelian group which is generated by any neighborhood of zero and let X be a complex linear space. If an exponential polynomial $f: G\to X$ vanishes on a measurable set of positive measure, then it vanishes everywhere.*

PROOF. Using the notations of the preceding theorem we prove the statement by induction on n, and for $n=1$ it follows from theorem 2.

Suppose that it is proved for $n\le k$ and let $n=k+1$. We assume that f vanishes on the compact set $K\subset G$ of positive measure. Similarly to the proof of the preceding theorem we have that $\varphi(x)=0$ for all x in K, and let $V\subset G$ be a neighborhood of zero for which $\lambda(K\cap K-y\cap\dots\cap K-(N+1)y)>0$ whenever y is in V. By the same argument as in theorem 3 there exists an element y in V for which $m_1(y)\neq m_2(y)$. Let $K_1=K\cap K-y\cap\dots\cap K-(N+1)y$, then $x+ky$ belongs to K for all x in K_1 $(k=0,1,\dots,N+1)$ and

$$\Delta_y^{N+1}\varphi(x)=\Delta_y^{N+1}p_1(x)+\sum_{i=2}^{k+1} m_i(x)m_1(x)^{-1}\left[(m_i(y)m_1(y)^{-1}-1)^{N+1}A_i^{(\mu_i)}(x)+q_{i,y}(x)\right]=0$$

for all x in K_1, where $q_{i,y}: G\to X$ is a polynomial of degree at most $\mu_i - 1$ $(i=2,\dots,k+1)$. Hence we have

$$\sum_{i=2}^{k+1} m_i(x)\left[\left(m_i(y)m_1(y)^{-1}-1\right)^{N+1} A_i^{(\mu_i)}(x) + q_{i,y}(x)\right] = 0$$

for all x in K_1 and by induction $\lambda K_1 > 0$ implies

$\left(m_2(y)m_1(y)^{-1}-1\right)^{N+1} A_2^{(\mu_2)}(x) + q_{2,y}(x) = 0$ for all x in G. Here

$m_2(y)m_1(y)^{-1} \neq 1$ and the degree of $q_{2,y}$ is at most $\mu_2 - 1$, hence $A_2^{(\mu_2)} = 0$, and the theorem is proved.

COROLLARY 5. Let G be a locally compact Abelian group which is generated by any neighborhood of the zero and let X be a complex topological linear space. Then the set of all zeros of a nonzero continuous exponential polynomial on G with values in X has measure zero.

REFERENCES

[1] Djoković, D.Z., A representation theorem for $(X_1-1)\dots(X_n-1)$ and its applications, Ann. Polon. Math. 22(1969), 189-198.

[2] Edgar, G.A., Rosenblatt, J.M. Difference equations over locally compact Abelian groups, Trans. Amer. Math. Soc. 253 (1979), 257-289.

[3] Federer, H., Geometric Measure Theory, Springer-Verlag, Berlin, Heidelberg, New York, 1969.

László Székelyhidi
Department of Mathematics
University of Debrecen
H-4010 Debrecen Pf. 12.

Received March 15, 1982

C.R. Math. Rep. Acad. Sci. Canada - Vol. IV, No. 4, August 1982 Août

THE GENERAL SYMMETRIC SOLUTION OF A FUNCTIONAL EQUATION ARISING IN THE MIXED THEORY OF INFORMATION

B. R. Ebanks

Presented by J. Aczél, F.R.S.C.

Abstract. We find all symmetric solutions of a functional equation holding for disjoint nonempty sets in a ring of sets. The general form of all symmetric, branching measures of information both in the mixed theory of information and on rings of sets are consequences.

1. Let B be a ring of subsets of a given set X, containing, with any two sets, also their union and difference, thus also their intersection and the empty set 0. (See [5].) In the determination of all inset information functions of degree α (see Aczél [1]), the functional equation

$$(1) \qquad \Delta(x \cup y, z) + \Delta(x,y) = \Delta(x \cup z, y) + \Delta(x,z)$$

arises, for pairwise disjoint $x,y,z \in B$. Because of considerations involved in the study of information measures of higher dimensions, Aczel has noted that it would be interesting to investigate inset information functions of degree α on the open domain (i.e. with both empty sets and 0 probabilities excluded). In our context, that means that x,y, and z must be nonempty. We shall find all symmetric solutions of (1) on D_3, where (for n = 1,2,...)

$$D_n = \{(x_1,x_2,\ldots,x_n) \mid 0 \neq x_i \in B,\ x_i \cap x_j = 0 \text{ for } i \neq j;\ i,j = 1,2,\ldots,n\}.$$

Theorem. *A map $\Delta\colon D_2 \to R$ (R the set of reals) satisfies*

$$(2) \qquad \Delta(x,y) + \Delta(x \cup y, z) = \Delta(x, y \cup z) + \Delta(y,z),$$

$$(3) \qquad \Delta(u,v) = \Delta(v,u)$$

for all $(x,y,z) \in D_3$ *and* $(u,v) \in D_2$ *if and only if there exists a* *map* $f: D_1 \to R$ *such that*

$$\Delta(x,y) = f(x) + f(y) - f(x \cup y),\ (x,y) \in D_2. \tag{4}$$

For any (3) symmetric Δ, (1) is equivalent to (2). Thus the theorem states that (4) is the general symmetric solution of (1). The proof will be given in section 2. Our first objective is to extend Δ to the algebra generated by B. Let $\Omega = \cup\{x \mid x \in B\} \subset X$ and $B' = \{x' \mid x \in B\}$, where $x' = \Omega \setminus x$. The following two observations are made without proof.

Lemma 1. $A = B \cup B'$ *is an algebra of subsets of* Ω. *Moreover, for all* $x,y \in A$, $x \cup y \in B$ *if* $x,y \in B$, *otherwise* $x \cup y \in B'$.

Lemma 2. *If* $\Omega \in B$, *then* $A = B$. *Otherwise*, $B \cap B' = 0$, *and at most one of* x,y *is in* B' *if* $x \cap y = 0$ $(x,y \in A)$.

Now, let $S_n = \{(x_1,x_2,\ldots,x_n) \mid 0 \neq x_i \in A,\ x_i \cap x_j = 0$ for $i \neq j;\ i,j = 1,2,\ldots,n\}$, $n = 1,2,\ldots$, and define $\overline{\Delta}$ on S_2 by

$$\overline{\Delta}(x,y) = \overline{\Delta}(y,x) = \begin{cases} \Delta(x,y) & \text{, if } x,y \in B \\ 0 & \text{, if } x \notin B,\ x \cup y = \Omega \\ \Delta(y,(x \cup y)') & \text{, if } x \notin B,\ x \cup y \subsetneqq \Omega \end{cases} . \tag{5}$$

Lemma 3. $\overline{\Delta} : S_2 \to R$ *defined by* (5) *satisfies*

$$\overline{\Delta}(x,y) + \overline{\Delta}(x \cup y,z) = \overline{\Delta}(x,y \cup z) + \overline{\Delta}(y,z), \tag{2'}$$

$$\overline{\Delta}(u,v) = \overline{\Delta}(v,u) \tag{3'}$$

for all $(x,y,z) \in S_3$ *and* $(u,v) \in S_2$.

Proof: (3') follows immediately from (5). We proceed to prove (2'). If $\Omega \in B$, then $(x,y,z) \in D_3$ by Lemma 2, and (2') is just (2). Suppose $\Omega \notin B$. If $(x,y,z) \in D_3$, then (2') is just (2), so suppose that at least one of x,y,z is in B'. By Lemma 2, exactly one of x,y,z is in B'.

Case 1. Suppose $z \in B'$. If $x \cup y \cup z = \Omega$, then by (5), (3), and Lemma 1,

$\bar{\Delta}(x,y) + \bar{\Delta}(x \cup y,z) = \Delta(x,y) = \Delta(y,x) = \Delta(y,(z \cup y)') = \bar{\Delta}(y,z) + \bar{\Delta}(x,y \cup z)$.

If $x \cup y \cup z \subsetneq \Omega$, then by (5), (3), (2), and Lemma 1,

$\bar{\Delta}(x,y) + \bar{\Delta}(x \cup y,z) = \Delta(x,y) + \Delta(x \cup y,(x \cup y \cup z)')$

$= \Delta(y,x) + \Delta(y \cup x,(y \cup x \cup z)') = \Delta(y,x \cup (y \cup x \cup z)') + \Delta(x,(y \cup x \cup z)')$

$= \Delta(y,(y \cup z)') + \Delta(x,(x \cup y \cup z)') = \bar{\Delta}(y,z) + \bar{\Delta}(x,y \cup z)$.

(Note that x,y,z are disjoint by definition of S_3.)

Cases 2, 3. Suppose $y \in B'$ or $x \in B'$. The proofs in these cases are similar to that in case 1.

2. Proof of Theorem: By Lemma 3, Δ can be extended to a map $\bar{\Delta}$ on S_2 satisfying (2') and (3'). Now, define a real-valued map K on all finite partitions $\{x_1,\dots,x_n\}$ of Ω into disjoint nonempty events $x_i \in A$ by $K(\{\Omega\})$ = an arbitrary real number and

(6) $$K(\{x_1,\dots,x_n\}) = K(\{x_1 \cup x_2,x_3,\dots,x_n\}) + \bar{\Delta}(x_1,x_2)$$

for all $(x_1,\dots,x_n) \in S_n$ such that $\cup x_i = \Omega$ $(n = 2,3,\dots)$. (2') and (3') guarantee exactly that K is well-defined, i.e. that

$$K(\{x_1,\dots,x_n\}) = K(\{x_{p(1)},\dots,x_{p(n)}\})$$

for all $(x_1,\dots,x_n) \in S_n$ and all permutations p on $\{1,..,n\}$, $n = 2,3,\dots$. By a result of Davidson and Ng [2], then

(7) $$K(\{x_1,\dots,x_n\}) = \sum_{i=1}^{n} f(x_i)$$

in terms of a map $f: A \setminus \{0\} \to R$. By (5), (6), (7), we have (4).

Corollary 1. A map $K: D_n \to R$ $(n = 1,2,3,\dots)$ satisfies

$$K(x_1,\dots,x_n) = K(x_1 \cup x_2,x_3,\dots,x_n) + \Delta(x_1,x_2)$$

and, for all permutations p on $\{1,2,\dots,n\}$,

$$K(x_1,\dots,x_n) = K(x_{p(1)},\dots,x_{p(n)})$$

<u>if and only if there exists a map</u> $f: D_1 \to R$ <u>such that</u>

$$K(x_1,\dots,x_n) = \sum_{i=1}^{n} f(x_i) + K\left(\bigcup_{i=1}^{n} x_i\right) - f\left(\bigcup_{i=1}^{n} x_i\right).$$

3. Our theorem also applies to the characterization of branching, symmetric inset entropies. Here, however, we do not restrict ourselves to the open domain, but consider entropies on $\overline{D}_n \times \Gamma_n$, where $\overline{D}_n = \{(x_1,\dots,x_n) \mid x_i \in B,\ x_i \cap x_j = 0 \text{ for } i \neq j;\ i,j = 1,2,\dots,n\}$

and $\Gamma_n = \{(p_1,\dots,p_n) \mid \Sigma p_i = 1;\ p_i \geq 0,\ i = 1,2\dots,n\}$.

<u>Corollary 2</u>. <u>A map</u> $I: \overline{D}_n \times \Gamma_n \to R$ $(n = 1,2,3,\dots)$ <u>is branching</u>,

$$I\begin{pmatrix} x_1,\dots,x_n \\ p_1,\dots,p_n \end{pmatrix} = I\begin{pmatrix} x_1 \cup x_2, x_3,\dots,x_n \\ p_1 + p_2, p_3,\dots,p_n \end{pmatrix} + \Delta\begin{pmatrix} x_1,x_2 \\ p_1,p_2 \end{pmatrix},$$

<u>and symmetric</u>, (<u>for all permutations</u> π <u>on</u> $\{1,\dots,n\}$)

$$I\begin{pmatrix} x_1,\dots,x_n \\ p_1,\dots,p_n \end{pmatrix} = I\begin{pmatrix} x_{\pi(1)},\dots,x_{\pi(n)} \\ p_{\pi(1)},\dots,p_{\pi(n)} \end{pmatrix},$$

<u>if and only if there exists a map</u> $\mu: B \times [0,1] \to R$ <u>such that, with</u> $\gamma(x) = I\binom{x}{1} - \mu(x,1)$,

$$(8)\quad I\begin{pmatrix} x_1,\dots,x_n \\ p_1,\dots,p_n \end{pmatrix} = \sum_{i=1}^{n} \mu(x_i,p_i) + \gamma\left(\bigcup_{i=1}^{n} x_i\right).$$

<u>Proof:</u> Symmetry and branching imply that Δ satisfies

$$(9)\quad \Delta\begin{pmatrix} x_1,x_2 \\ p_1,p_2 \end{pmatrix} = \Delta\begin{pmatrix} x_2,x_1 \\ p_2,p_1 \end{pmatrix},$$

$$(10)\quad \Delta\begin{pmatrix} x_1,x_2 \\ p_1,p_2 \end{pmatrix} + \Delta\begin{pmatrix} x_1 \cup x_2, x_3 \\ p_1 + p_2, p_3 \end{pmatrix} = \Delta\begin{pmatrix} x_1, x_2 \cup x_3 \\ p_1, p_2 + p_3 \end{pmatrix} + \Delta\begin{pmatrix} x_2,x_3 \\ p_2,p_3 \end{pmatrix}$$

for all $(x_1,x_2,x_3) \in \overline{D}_3$ and $p_1,p_2,p_3 \geq 0$ with $p_1 + p_2 + p_3 \leq 1$. By Lemma 4.2 in [4], (10) has general solution

$$(11)\quad \Delta\binom{x,y}{p,q} = \eta(x,y) + \tilde{\mu}(x,p) + \tilde{\mu}(y,q) - \tilde{\mu}(x \cup y, p + q) + \tilde{\psi}\binom{x,y}{p,q},$$

where $\eta: \bar{D}_2 \to R$ satisfies (2) and $\tilde{\psi}$ is antisymmetric and bi-additive, i.e.

(12) $$\tilde{\psi}\begin{pmatrix} x,y \\ p,q \end{pmatrix} = -\tilde{\psi}\begin{pmatrix} y,x \\ q,p \end{pmatrix}, \quad \tilde{\psi}\begin{pmatrix} x\cup y,z \\ p+q,r \end{pmatrix} = \tilde{\psi}\begin{pmatrix} x,z \\ p,r \end{pmatrix} + \tilde{\psi}\begin{pmatrix} y,z \\ q,r \end{pmatrix}.$$

The purpose of what follows is to essentially eliminate the η and $\tilde{\psi}$ terms. Divide η into symmetric and antisymmetric parts

(13) $$S(x,y) = \tfrac{1}{2}\Big[\eta(x,y) + \eta(y,x)\Big], \quad A(x,y) = \tfrac{1}{2}\Big[\eta(x,y) - \eta(y,x)\Big],$$

so that $\eta = S + A$. Since S satisfies (2) and (3), our Theorem gives

(14) $$S(x,y) = f(x) + f(y) - f(x\cup y),$$

as long as x and y are nonempty. But with $y = 0$, (2) for S yields $S(x,0) = S(0,z)$ for all disjoint $x,z \in B$. Thus, defining $f(0)$ to be this common value, we have (14) on $\bar{D}_2$.

Furthermore, A is antisymmetric, hence the map ψ defined by

(15) $$\psi\begin{pmatrix} x,y \\ p,q \end{pmatrix} = \tilde{\psi}\begin{pmatrix} x,y \\ p,q \end{pmatrix} + A(x,y)$$

is also antisymmetric. Now (11), (13), (14), and (15) give

$$\Delta\begin{pmatrix} x,y \\ p,q \end{pmatrix} = \mu(x,p) + \mu(y,q) - \mu(x\cup y,p+q) + \psi\begin{pmatrix} x,y \\ p,q \end{pmatrix},$$

where $\mu(x,p) = \tilde{\mu}(x,p) + f(x)$. But the (9) symmetry of Δ implies that ψ is also symmetric. Thus $\psi \equiv 0$, and

$$\Delta\begin{pmatrix} x,y \\ p,q \end{pmatrix} = \mu(x,p) + \mu(y,q) - \mu(x\cup y,p+q).$$

Therefore, by the branching property, I has representation (8) with $\gamma(x) = I\binom{x}{1} - \mu(x,1)$. The converse is easily checked.

Remarks. Corollary 2 is an improvement on Theorem 1.1 in [3]. (See Remark 1.2 in [3].) A desirable further improvement would be a corresponding result for branching, symmetric maps on the open domain $D_n \times \overset{\circ}{\Gamma}_n$, where $\overset{\circ}{\Gamma}_n = \{(p_1,\dots,p_n) \mid \Sigma p_i = 1;\ p_i > 0,\ i = 1,2,\dots,n\}$.

Acknowledgment. The author would like to express his sincere thanks to C.T. Ng for several helpful suggestions regarding an earlier version of this paper.

REFERENCES

[1] Aczél, J., A mixed theory of information. VII. Inset information functions of all degrees, C. R. Math. Rep. Acad. Sci. Canada 2 (1980), 125-129.

[2] Davidson, K. R.; Ng, C. T., Information measures and cohomology, Utilitas Math. 20 (1981), 27-34.

[3] Ebanks, B. R., Branching and generalized-recursive inset entropies, Proc. Amer. Math. Soc. 79 (1980), 260-267.

[4] Ebanks, B. R., A characterization of separable utility functions, Kybernetika 17 (1981), 244-255.

[5] Halmos, P. R., Measure Theory. Van Nostrand, Toronto, New York, London, 1950, pp. 19-22.

Department of Math., Texas Tech University, Lubbock, Texas 79409
U.S.A.

Received April 4, 1982

C.R. Math. Rep. Acad. Sci. Canada - Vol. IV, No. 4, August 1982 Août

ASYMPTOTIC ESTIMATION OF THE COEFFICIENTS OF THE CONTINUED FRACTION REPRESENTING THE BINET FUNCTION[1]

J. Cizek and E.R. Vrscay

Presented by G.F.D. Duff, F.R.S.C.

Abstract: The continued fraction for the Binet function is studied with the use of the quotient-difference formalism. Two conjectures based on extensive numerical evidence are formulated. The most important result is that for large n the coefficients in the continued fraction expansion of the Binet function appear to be asymptotically equal to $\frac{1}{16}n^2$.

The Binet function which plays an important role both in mathematics and theoretical physics is defined as [1]

$$J(z) = \log \Gamma(z) - (z - \tfrac{1}{2}) \log z + z - \tfrac{1}{2} \log 2\pi \;. \qquad (1)$$

There exist several integral expressions for this function and its asymptotic series in powers of z^{-1} is in all respects very well known [2]. Knowledge of the continued fraction expansion for $J(z)$ is, on the other hand, very scarce. In fact, until very recently only seven coefficients in the expansion[2]

$$J(z) = \frac{a_1|}{|z} + \frac{a_2|}{|z} + \dots \qquad (2)$$

were known [2,3]. Now Char [4] has calculated the first 41 coefficients by means of the algebraic manipulation system MACSYMA. The structure of coefficients in

[1] This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada.

[2] The indexing of the continued fraction coefficients a_n in this paper follows the scheme of References 1 and 3, where a_1 denotes the first coefficient of expansion (2).

(2) seems to be very complicated, as already mentioned by Stieltjes in his classic paper [5] and confirmed by Henrici [1].

In order to elucidate the nature of the a_n it is very instructive to study the function

$$K(z) = \frac{1}{3}\int_0^\infty \frac{e^{-4zt}}{\cosh t}\,dt \tag{3}$$

which is closely related to $J(z)$. The continued fraction representation for $K(z)$ is well known [6],

$$K(z) = \frac{b_1|}{|z} + \frac{b_2|}{|z} + \frac{b_3|}{|z} + \dots \quad , \tag{4}$$

where $b_1 = \frac{1}{12}$ and for $n \geq 1$

$$b_{n+1} = \frac{1}{16}\,n^2 \; . \tag{5}$$

In order to study the relation of $J(z)$ and $K(z)$ let us introduce the following two functions and their asymptotic expansions

$$f(u) = \frac{1}{\sqrt{u}}\,J(\sqrt{u}) = \sum_{n=0}^{\infty} \frac{c_n}{u^{n+1}} \tag{6}$$

$$g(u) = \frac{1}{\sqrt{u}}\,K(\sqrt{u}) = \sum_{n=0}^{\infty} \frac{d_n}{u^{n+1}} \tag{7}$$

From (2) and (4) the continued fraction expansions for $f(u)$ and $g(u)$ may be given as

$$f(u) = \frac{a_1|}{|u} + \frac{a_2|}{|1} + \frac{a_3|}{|u} + \dots \tag{8}$$

$$g(u) = \frac{b_1|}{|u} + \frac{b_2|}{|1} + \frac{b_3|}{|u} + \dots \quad . \tag{9}$$

The series coefficients c_n and d_n in (6) and (7) are given by

$$|c_n| = \frac{|B_{2n+2}|}{(2n+2)(2n+1)} \quad , \quad |d_n| = \frac{1}{12}\frac{|E_{2n}|}{4^{2n}} \quad , \tag{10}$$

where B_j and E_j represent the Bernoulli and Euler numbers respectively [3]. The asymptotic expressions for the series coefficients c_n and d_n are thus dominated by the term $(2n)!$.

We shall be using the quotient-difference (QD) formalism [1] to learn about the properties of $f(u)$ from $g(u)$. The QD scheme associates a unique Stieltjes-type continued fraction of the form (8) or (9) to a given formal power series, subject to some restrictions on the series. The elements of the relevant QD schemes will be denoted as follows:

f(u) QD scheme:			g(u) QD scheme:		
q_1^0			r_1^0		
	e_1^0			f_1^0	
q_1^1		q_2^0	r_1^1		r_2^0
	e_1^1	$\vdots \ddots$		f_1^1	$\vdots \ddots$
q_1^2	$\vdots$		r_1^2	$\vdots$	
$\vdots$			$\vdots$		

(11)

The first elements of each column in each QD array, $\{q_1^0, e_1^0, q_2^0, \ldots\}$ and $\{r_1^0, f_1^0, r_2^0, \ldots\}$, form the sequences of coefficients a_n and b_n of the continued fraction representations in (8) and (9). The QD arrays are built up in a triangular fashion from their first columns, as illustrated in (11), by means of a set of well-defined "rhombus rules" [1]. The elements ("initial values") in the first columns are given by

$$q_1^n = \frac{|c_{n+1}|}{|c_n|} \quad , \qquad r_1^n = \frac{|d_{n+1}|}{|d_n|} \quad . \tag{12}$$

From the well-known expressions for the Bernoulli and Euler numbers in terms of factorial and, respectively, Riemann and Lerch functions [3] it can easily be shown that

$$\lim_{n\to\infty} y_n = 1 \quad , \tag{13}$$

where

$$y_n = q_1^n / r_1^n \quad . \tag{14}$$

Having in mind the properties of the Riemann and Lerch functions it is not surprising that the ratio y_n is converging to the value 1 very quickly. To illustrate this rapid convergence, the following values are given here: $y_0 = 0.533333$, $y_1 = 0.914286$, $y_2 = 0.983607$, $y_5 = 0.999810$, $y_{10} = 1.000000$. This means that the first column of the QD scheme for $f(u)$ is very closely related to the scheme for $g(u)$. Since the first column of entries completely determines the entire QD array we can expect the arrays corresponding to $f(u)$ and $g(u)$ to be quite closely related. In fact, these two QD schemes appear to be "interwoven". This interesting and important property is formulated in

<u>Conjecture 1:</u> $|q_k^n| < |r_k^n|$ and $|e_k^n| > |f_k^n|$, $n \geq 0$, $k \geq 1$. (15)

This conjecture was verified by numerical calculations of the QD triangles (IBM quadruple precision) which permitted the calculation of the continued fraction coefficients a_{36} and b_{36} respectively. We have the observed set of relations:

$$\begin{array}{rcccccl}
0.033333 = q_1^0 & = a_2 & < & b_2 & = & r_1^0 & = 0.062500 \\
0.253303 = e_1^0 & = a_3 & > & b_3 & = & f_1^0 & = 0.250000 \\
0.525606 = q_2^0 & = a_4 & < & b_4 & = & r_2^0 & = 0.562500 \\
 & \vdots & & \vdots & & & \\
72.355941 = e_{17}^0 & = a_{35} & > & b_{35} & = & f_{17}^0 & = 72.250000 \\
76.424655 = q_{18}^0 & = a_{36} & < & b_{36} & = & r_{18}^0 & = 76.562500 \quad .
\end{array} \tag{16}$$

The a_n sequence can be seen to oscillate around the b_n sequence defined expressly by (5). Numerical analysis of the differences between a_n and b_n suggests that the asymptotic behaviour of the a_n is as follows:

<u>Conjecture 2</u>: $a_n = b_n + R_n$ (17)

where R_n is of the order n .

This asymptotic information is very useful both for theoretical and practical reasons. There exist different methods for the estimation of the convergence of a continued fraction. Henrici has presented three such methods [1]. With a knowledge of the coefficients a_n of the continued fraction, a much better estimate may be made as compared to that estimate obtained from the methods based on the knowledge of the series coefficients c_n or d_n , respectively. We have found such an estimate but have postponed its publication until exact proofs of Conjectures 1 and 2 are presented.

Let us finally mention that we are also studying other power series expansions whose coefficients a_n are dominated by the term $(2n)!$. Series of this type are of special interest because by Carleman's condition [1,2] their continued fraction representations "lie on the borderline of convergence".

References

[1] P. Henrici: Applied and Computational Complex Analysis, Volumes I and II, Wiley (1977).

[2] W.B. Jones and W.J. Thron: Continued Fractions, Analytic Theory and Applications, Addison-Wesley (1980), 348-350.

[3] M. Abramowitz and I. Stegun: Handbook of Mathematical Functions, Dover (1965).

[4] B.W. Char: On Stieltjes' Continued Fraction for the Gamma Function, Math. of Computation 34 (1980), 547-551.

[5] T.J. Stieltjes: Recherches sur les fractions continues, Ann. Fac. Sci. Toulouse 8 (1894), 1-122.

[6] T.J. Stieltjes: Sur la réduction en fraction continue d'une série procédant suivant les puissances descendantes d'une variable, Ann. Fac. Sci. Toulouse 3 (1889), 1-17.

Department of Applied Mathematics
Faculty of Mathematics
University of Waterloo
Waterloo, Ontario N2L 3G1

Received April 23, 1982

C.R. Math. Rep. Acad. Sci. Canada - Vol. IV, No. 4, August 1982 Août

INFORMATION FUNCTIONS ON OPEN DOMAIN IV

Pl. Kannappan*

Presented by J. Aczél, F.R.S.C.

Abstract. We consider functional equations on open domain similar to the fundamental equation of information treated in [4,1,3]. We determine the general solution of the multiplace functional equation

(1) $$f(x,y)+(1-x)^{\beta}g\left(\frac{u}{1-x},\frac{v}{1-y}\right)=h(u,v)+(1-u)^{\beta}k\left(\frac{x}{1-u},\frac{y}{1-v}\right) \quad \text{on} \quad D^{n+1}$$

for $(x,u)\in D$, $(y,v)\in D^n$ (β is a fixed real), where

(2) $$D=\{(r,s)\,|\,r,s,r+s\in I=]0,1[\},$$

without any assumptions on the real valued functions f,g,h and k.

THEOREM. The general solutions $f,g,h,k\colon I^{n+1}\to\mathbb{R}$(reals)$(x\in I,\ y\in I^n)$ of equation (1) for $(x,u)\in D$, $(y,v)\in D^n$ are given by

(i) $$\begin{cases} f(x,y)=\ell_1(1-x)+\ell_2(x)+M_1(1-y)+M_2(y)+A \\ g(x,y)=\ell_1(1-x)-\ell_3(1-x)+\ell_3(x)+M_1(1-y)-M_3(1-y)+M_3(y)+B-A+C \\ h(x,y)=\ell_1(1-x)+\ell_2(1-x)-\ell_3(1-x)+\ell_3(x)+M_1(1-y)+M_2(1-y) \\ \qquad\qquad -M_3(1-y)+M_3(y)+B \\ k(x,y)=\ell_1(1-x)-\ell_3(1-x)+\ell_2(x)+M_1(1-y)-M_3(1-y)+M_2(y)+C \end{cases} \quad \text{if } \beta=0;$$

(ii) $$\begin{cases} f(x,y)=S(x,y)+ax+c_1+c_2-a \\ g(x,y)=S(x,y)+(b_1+b_2-c_1-c_2)x+a-b_1 \\ h(x,y)=S(x,y)+(b_2-c_1)x+c_1 \\ k(x,y)=S(x,y)+b_1x+c_2-b_1, \end{cases} \quad \text{if } \beta=1;$$

---

*Research partially supported by a NSERC of Canada grant.

$$\text{(iii)}\quad \begin{cases} f(x,y) = \epsilon(\beta)d(x)+ax^{\beta}+a_2(1-x)^{\beta}+a_3 \\ g(x,y) = -\epsilon(\beta)d(x)+b_1x^{\beta}+b_2(1-x)^{\beta}-a_2 \\ h(x,y) = -\epsilon(\beta)d(x)+b_1x^{\beta}+b_3(1-x)^{\beta}+a_3 \\ k(x,y) = \epsilon(\beta)d(x)+a_1x^{\beta}+b_2(1-x)^{\beta}-b_3 \end{cases} \quad \text{if } \beta\neq 0,1.$$

Here $\epsilon(2)=1$, $\epsilon(\beta)=0$ if $\beta\neq 2$, A,B,C,a,a_i,b_i,c_i are constants and ℓ_i, M_i, d satisfy

$$(3)\qquad \ell(xu) = \ell(x)+\ell(u) \qquad (x,u\in\,]0,1[)$$

$$(4)\qquad M(yv) = M(y)+M(v) \qquad (y,v\in I^n)$$

$$(5)\qquad \begin{cases} d(x+u) = d(x)+d(u) \\ d(xu) = xd(u)+ud(x) \quad (x,u\in\mathbb{R})\ \text{(a real derivation)} \end{cases}$$

respectively, and $S(x,y) = x\ell_1(x)+(1-x)\ell_1(1-x)+xM_1(y)+(1-x)M_1(1-y)$.

<u>Remark 1</u>. In the proof we make use of the solution [2] of

$$(6)\qquad F(x)+(1-x)^{\beta}G\left(\frac{u}{1-x}\right) = H(u)+(1-u)^{\beta}K\left(\frac{x}{1-u}\right) \quad \text{on } I$$

and of the linear independence of $\ell_1(x)$, $\ell_2(1-x)$, 1, of $x\ell(x)$, $(1-x)\ell(1-x)$, x, 1 and of 1, x^{β}, $(1-x)\ell(1-x)$ $(\beta\neq 0,1)$ $(\ell_1,\ell_2,\ell\not\equiv 0$ satisfy (3)).

<u>Proof</u>: <u>Case 1</u>: $\beta=0$. Then (1) becomes

$$(7)\qquad f(x,y)+g\left(\frac{u}{1-x},\frac{v}{1-y}\right) = h(u,v)+k\left(\frac{x}{1-u},\frac{y}{1-v}\right) \quad \text{on } D^{n+1}.$$

The proof is by induction. Solution (i) can be rewritten as

$$(i^1)\quad \begin{cases} f(\xi) = L(1-\xi)+M(\xi)+A \\ g(\xi) = L(1-\xi)-N(1-\xi)+N(\xi)+B-A+C \\ h(\xi) = L(1-\xi)+M(1-\xi)-N(1-\xi)+N(\xi)+B \\ k(\xi) = L(1-\xi)-N(1-\xi)+M(\xi)+C \end{cases}$$

for $\xi=(x,y)\in I^{n+1}$ where L,M,N are solutions of (4) on I^{n+1} and A,B,C are arbitrary constants. For fixed $y,v\in I^n$, (7) becomes, with $F(x)=f(x,y)$, $G(x)=g(x,\frac{y}{1-v})$, $H(x)=h(x,v)$, $K(x)=k(x,\frac{v}{1-y})$,

$$F(x)+G(\frac{u}{1-x})=H(u)+K(\frac{x}{1-u})$$

i.e. (6) with $\beta=0$, so that [2] gives for $x\in I$,

$$(8)\qquad \begin{cases} F(x)=f(x,y)=\ell_1(1-x)+\ell_2(x)+a \\ G(x)=g(x,\frac{y}{1-v})=\ell_1(1-x)-\ell_2(1-x)+\ell_3(x)+b-a+c \\ H(x)=h(x,v)=\ell_1(1-x)+\ell_2(1-x)-\ell_3(1-x)+\ell_3(x)+b \\ K(x)=k(x,\frac{y}{1-v})=k(x,\frac{1}{1-v})=\ell_1(1-x)+\ell_3(1-x)+\ell_2(x)+c \end{cases}$$

where the solutions ℓ_i of (3) and the 'constants' a,b,c may depend upon y and v. Using the linear independence of $\ell_1(1-x)$, $\ell_2(x)$, 1 (Remark 1), from the forms of f,g,h,k given by (8) we have

$$(9)\qquad \ell_2(x,y,v)=m_2(x,y),\quad a(y,v)=a(y)$$

$$(10)\qquad \begin{cases} (\ell_1-\ell_3)(1-x,y,v)=m_3(1-x,\frac{v}{1-y}), \\ \ell_3(x,y,v)=m_4(x,\frac{v}{1-y}),\quad (b-a+c)(y,v)=e(\frac{v}{1-y}) \end{cases}$$

$$(11)\qquad \ell_3(x,y,v)=m_5(x,v),\quad b(y,v)=b(v)$$

$$(12)\qquad \begin{cases} (\ell_1-\ell_3)(1-x,y,v)=m_6(1-x,\frac{y}{1-v}), \\ \ell_2(x,y,v)=m_7(x,\frac{y}{1-v}),\quad c(y,v)=c(\frac{y}{1-v}) \end{cases}$$

where the m_i's for fixed y,v are solutions of (3).

From (9) and (12), $m_2(x,y) = m_7(x,\frac{y}{1-v})$. The left side is independent of v, so m_2 and m_7 are independent of the second variable, i.e. functions of x only. Hence $\ell_2(x,y,v)$ depends only upon x. Similarly, from (10), (11) ℓ_3 and from (10), (12) $\ell_1 - \ell_3$, thus also ℓ_1 depend only upon x. Now, from (9) to (12),

$$b(v)+c(\frac{y}{1-v}) = a(y)+e(\frac{v}{1-y}) \quad \text{on } D^n. \tag{13}$$

Then by the induction hypothesis (i^1), we have

$$\begin{cases} a(y) = M_1(1-y)+M_2(y)+A \\ e(y) = M_1(1-y)-M_3(1-y)+M_3(y)+B-A+E \\ b(y) = M_1(1-y)+M_2(1-y)-M_3(1-y)+M_3(y)+B \\ c(y) = M_1(1-y)-M_3(1-y)+M_2(y)+E \end{cases} \tag{14}$$

on I^n, where the M_i's satisfy (4) and A,B,E are constants so that f,g,h,k indeed have the form given by (i).

<u>Case 2</u>. $\beta = 1$. For fixed $y,v \in I^n$, (1) is of the form (6). From [2],

$$\begin{cases} F(x) = f(x,y) = x\ell(x)+(1-x)\ell(1-x)+c_1x+c_2 \\ G(x) = g(x,\frac{v}{1-y}) = x\ell(x)+(1-x)\ell(1-x)+c_3x+c_4 \\ H(x) = h(x,v) = x\ell(x)+(1-x)\ell(1-x)+c_5x+c_4+c_2-c_5+c_3 \\ K(x) = k(x,\frac{y}{1-v}) = x\ell(x)+(1-x)\ell(1-x)+(c_1-c_4)x+c_5-c_3 \end{cases} \tag{15}$$

where ℓ and c_i may depend upon y and v.

Using the linear independence of $x\ell(x)$, $(1-x)\ell(1-x)$, x and 1 (Remark 1), (15) shows that ℓ is independent of y and v and then

$$(16)\quad \begin{cases} c_1(y,v) = c_1(y)\,(\text{say}), \quad c_2(y,v) = c_2(y), \\ c_3(y,v) = c_3(\frac{v}{1-y}), \quad c_4(y,v) = c_4(\frac{v}{1-y}), \quad c(y,v) = c_5(v), \end{cases}$$

$$(17)\quad c_1(y) - c_4(\tfrac{v}{1-y}) = B'(\tfrac{y}{1-v})\,(\text{say}),$$

$$(18)\quad c_5(v) - c_3(\tfrac{v}{1-y}) = C'(\tfrac{y}{1-v})\,(\text{say}),$$

$$(19)\quad (c_4+c_3)(\tfrac{v}{1-y}) + c_2(y) - c_5(v) = A'(v)\,(\text{say}) \quad \text{on } D^n.$$

Now (17) is of the form (13). By (14), since $M_1 + M_2 = 0 = M_3$,

$$(20)\quad \begin{cases} c_1(y) = M_4(1-y) - M_4(y) + d_1, \quad c_4(t) = -M_4(1-t) + d_2 \\ B'(y) = M_4(1-y) - M_4(y) + d_3 \end{cases}$$

on I^n where M_4 (the old M_1) satisfies (4) and d_1, d_2, d_3 are constants. Since (18) and (19) are also of the form (13), by (14), we obtain

$$(21)\quad \begin{cases} c_5(v) = M_5(1-v) - M_5(v) + d_4, \quad c_3(t) = M_5(1-t) - M_5(t) + d_5 \\ C'(t) = -M_5(1-t) + d_6 \end{cases}$$

$$(22)\quad \begin{cases} c_2(y) = M_6(1-y) + d_7, \quad (c_3+c_4)(t) = M_6(t) + d_8 \\ (c_5+A')(v) = M_6(v) + d_9 \end{cases}$$

where M_5, M_6 satisfy (4) and the d's are constants.

Considering c_3, c_4 in (20), (22), we have $M_4 = M_5 = -M_6$. By (19), $A'(v) = -M_4(1-v) + d_9 - d_4$. So (15), (20), (21), (22) give (ii).

__Case 3__. $\beta \neq 0,1$. Take first __$\beta = 2$__. Here again, for fixed $y, v \in I^n$, (1) is a special case of (6) ($\beta = 2$), so that from [2], we get

$$(23)\quad \begin{cases} F(x) = f(x,y) = d(x) + a_1 x^\beta + a_2(1-x)^\beta + a_3 \\ G(x) = g(x,\frac{v}{1-y}) = -d(x) + b_1 x^\beta + a_4(1-x)^\beta - a_2 \\ H(x) = h(x,v) = -d(x) + b_1 x^\beta + b_2(1-x)^\beta + a_3 \\ K(x) = k(x,\frac{y}{1-v}) = d(x) + a_1 x^\beta + a_4(1-x)^\beta - b_2 \end{cases}$$

where d (satisfying (5)) and the 'constants' a_i, b_i depend upon y,v.

As before, using Remark 1, from (23) we conclude first that d is independent of y,v. The argument reducing a_1 (which occurs in f as a function of y and in k as a function of $y/1-v$) to a constant (independent of y,v) is the same as in case 1. Similarly, it can be shown that a_2,a_3,a_4,b_1,b_2 are also constants, proving (iii) if $\beta=2$. A similar argument that (iii) holds also if $\beta\neq 0,1,2$. □

Remark 2. Note that the solutions in the cases $\beta\neq 0,1$ depend only upon the first variable x. This is related to the special form of the factors $(1-x)^{\beta}$ and $(1-u)^{\beta}$ in the second and fourth terms of (1), which depend only on the variables in the first place. If the factors are of the form $\prod_{k=1}^{n+1}(1-x_k)^{\beta_k}$, etc. then the solutions may depend also upon the other variables.

I express my sincere thanks to Professor J. Aczél for his helpful remarks during the preparation of this paper.

References

1. J. Aczél, Information functions on open domain, III. C.R. Math. Rep. Acad. Acad. Sci. Canada, 2 (1980), 281-287.

2. Gy. Maksa, Solution on the open triangle of the generalized fundamental equations of information with four unknown functions. Utilitas Math.

3. Gy. Maksa and C.T. Ng, The fundamental equation of information on open domain. Publ. Math. Debrecen.

4. C.T. Ng, Information functions on open domains, I and II. C.R. Math. Rep. Acad. Sci. Canada 2 (1980), 119-123 and 155-158.

Department of Pure Mathematics
University of Waterloo
Waterloo, Ontario Canada N2L 3G1

Received May 1, 1982

C.R. Math. Rep. Acad. Sci. Canada - Vol. IV, No. 4, August 1982 Août

STRONGLY NONLINEAR VARIATIONAL INEQUALITIES

M. ASLAM NOOR

Presented by P. Ribenboim, F.R.S.C.

ABSTRACT:
The existence and uniqueness of the solution of a class of strongly nonlinear variational inequalities is considered. An iterative scheme is given to obtain the approximate solution of the variational inequalities.

Variational inequalities are now fundamental in the study of nonlinear problems having some extra constraint conditions in the study of fluid dynamics, plasma physics, chemical reactor theory and many other branches of mathematical and engineering sciences, see Duvaut and Lions[2], Glowinski, Lions and Tremoliers[3] and Noor[7]. In this paper, we consider the problem of strongly nonlinear variational inequalities in a Hilbert space. An iterative method is given to find the approximate solution of variational inequalities. It is also shown that the approximate solution obtained by the iterative scheme converges strongly in the Hilbert space to the exact solution.

To be more precise, let H be a real Hilbert space with its dual H', whose inner product and norm are denoted by $(.,.)$ and $\|.\|$ respectively. The pairing between $f\in H'$ and $u\in H$ is denoted by $\langle .,.\rangle$. Let M be a closed non-empty convex subset of H. We consider the problem of finding the minimum of the nonlinear functional $I[v]$, defined by

$$I[v] = L(v) - F(v), \qquad \text{for all } v\in H. \tag{1}$$

Many mathematical and engineering science problems either arise or can be reformulated in this form. Here one seeks to minimize the functional $I[v]$ over a whole space or on a convex set M in H. It is well known [3,6,7] that if F is a linear continuous functional on H, then the element $u\in M$ wchich minimizes $I[v]$ on M is given

$$\langle L'(u),v-u\rangle \geq \langle F,v-u\rangle, \qquad \text{for all } v\in M, \tag{2}$$

where $L'(u)$ is the Frechet differential of the nonlinear functional L at $u\in M$. For a Frechet differentiable nonlinear functional F, it was shown [10] that the minimum of $I[v]$ on M can be characterized by the inequality

$$\langle L'(u), v-u\rangle \geq \langle F'(u), v-u\rangle, \qquad \text{for all } v\in M. \quad (3)$$

Recently Toscano and Maceri [12] and Miersemann[5] have shown that the problem of elastic beams under unilateral constraints can be formulated in the form of inequality (3). Such type of inequalities are known as strongly non-linear variational inequalities. The main motivation of this paper is to show that under certain conditions there does exist a unique solution of a more general variational inequality of which (3) is a special case.

Let us consider the following problem:

<u>PROBLEM</u> 1. *Find* $u\in M$ *such that*

$$\langle Tu, v-u\rangle \geq \langle A(u), v-u\rangle, \qquad \text{for all } v\in M \quad (4)$$

where T and A are nonlinear operators.

For $M = H$, the problem 1 is equivalent to finding $u\in H$ such that

$$\langle Tu, v\rangle = \langle A(u), v\rangle, \qquad \text{for all } v\in H,$$

a case considered and studied by Noor[9] and Schechter, Shapiro and Snow[13] in a real Hilbert space and Banach spaces respectively.

Also note that if $A(u)$ is independent of u, that is $A(u)= f$ (say), then problem 1 is equivalent to finding $u\in M$ such that

$$\langle Tu, v-u\rangle \geq \langle f, v-u\rangle, \qquad \text{for all } v\in M.$$

This class of variational inequalities was considered and studied by Browder[1]. First of all, we define some notions.

<u>DEFINITIONS</u>. An operator $T:M \longrightarrow H'$ is called

(i) <u>Strongly Monotone</u>, if there exists a constant $\rho > 0$ such that

$$\langle Tu-Tv, u-v\rangle \geq \rho\|u-v\|^2, \qquad \text{for all } u,v\in M$$

(ii) <u>Antimonotone</u> , if

$$\langle Tu-Tv, u-v\rangle \leq 0, \qquad \text{for all } u,v\in M.$$

(iii) <u>Lipschitz continuous</u>, if there exists a constant $\mu > 0$ such that

$$\|Tu-Tv\| \leq \mu\|u-v\|, \qquad \text{for all } u,v\in M.$$

We also define Λ, a cannonical isomorphism from H' onto H, by

$$\langle f, v\rangle = (\Lambda f, v), \qquad \text{for all } f\epsilon H',\ v\epsilon H. \qquad (5)$$

We make the following hypothesis.

CONDITION N.

We assume that $\gamma<\rho$, where γ is the Lipschitz constant of the nonlinear operator A and ρ is the strongly monotonicity constant of T.

We now state and prove the main result.

THEOREM 1. Let T be a strongly monotone Lipschitz continous operator and A be a Lipschitz continuous antimonotone operator. If condition N holds, then there exists a unique solution $u\epsilon M$ such that (4) holds.

We need the following results.

LEMMA 1. Let ξ be a number such that $0<\xi< 2(\rho-\gamma)/(\mu^2-\gamma^2)$ and $\xi\gamma< 1$. Then there exists a θ with $0<\theta<1$ such that

$$\|\phi(u_1)-\phi(u_2)\| \leq \theta\|u_1-u_2\|, \qquad \text{for all } u_1,u_2\epsilon H,$$

where μ is the Lipschitz constant of T and for $u\epsilon H$, $\phi(u)\epsilon H'$ is defined by

$$\langle\phi(u),v\rangle = (u,v)-\xi\langle Tu,v\rangle + \xi\langle A(u),v\rangle \qquad \text{for all } v\epsilon H. \qquad (6)$$

PROOF: It can be proved by using the technique of Noor[8,11].

LEMMA 2: [6,8]. Let M be a convex subset of H. Then, given $z\epsilon H$, we have $x = P_M z$, if and only if

$$x\epsilon M : (x-z,y-x)\geq 0, \qquad \text{for all } y\epsilon M,$$

where P_M is a projection of H into M.

LEMMA 3: [6]. P_M is nonexpansive, i.e., $\|P_M u-P_M v\|\leq\|u-v\|$, for all $u,v\epsilon H$.

Using the technique of Lions-Stampacchia[4] and Noor[8], we now prove the theorem 1.

PROOF OF THEOREM 1.

(a). UNIQUENESS. see Noor [8].

(b) EXISTENCE. For a fixed ξ as in lemma 1 and $u\epsilon H$, define $\phi(u)\epsilon H'$ by (6). By lemma 2, there exists a unique $w\epsilon M$ such that

$$(w,v-w) \geq \langle\phi(u),v-w\rangle, \qquad \text{for all } v\epsilon M$$

and w is given by

$$w = P_M\Lambda\phi(u) = Tu$$

which defines a map from H into M.

Now for all $u_1, u_2 \varepsilon H$,

$$\|Tu_1 - Tu_2\| = \| P_M \Lambda\phi(u_1) - P_M \Lambda\phi(u_2)\| \leq \|\Lambda(u_1) - \Lambda(u_2)\|, \text{ see[6].}$$
$$\leq \|\phi(u_1) - \phi(u_2)\| \leq \theta \|u_1 - u_2\|, \text{ by lemma 1.}$$

Since $\theta < 1$. Tu is a contraction and has a fixed point Tu=u, which belongs to M, a closed convex set and satisfies

$$(u,v-u) \geq \langle\phi(u),v-u\rangle = (u,v-u) - \xi\langle Tu,v-u\rangle + \xi\langle A(u),v-u\rangle.$$

Thus for $\xi > 0$,

$$\langle Tu,v-u\rangle \geq \langle A(u),v-u\rangle, \qquad \text{for all } v \varepsilon M,$$

showing that u is a unique solution of problem 1.

REMARKS.

i. It is obvious that for Tu= L'(u), and A(u)=F'(u), the existence of a unique solution of a variational inequality (3) follows under the assumptions of theorem 1.

ii. If A is independent of u, that is A(u)=f, (say), then Lipschitz's constant γ is zero and lemma 1 reduces to a result of Noor[7] and ξ is a number $0 < \xi < 2\rho/\mu^2$. Consequently theorem 1 is exactly the same as one proved by Browder[1] and Noor[7].

iii. If T is a bilinear form and A(u) is independent of u, then theorem 1 reduces to the result of Lions and Stampacchia[4].

iv. For M = H, the problem 1 is equivalent to finding $u \varepsilon H$ such that

$$\langle Tu,v\rangle = \langle A(u),v\rangle, \qquad \text{for all } v \varepsilon H,$$

a problem considered by Noor[9] and its various special cases.

The problem of finding the approximate solution of variational inequalities is not an easy one. Due to the presence of the constraint, the approximate solution is no longer the projection of the exact solution as in the absence of constraints. In this section, we show that the approximate solution can be obtained by an iterative scheme. The convergence of the approximate solution to the exact solution is also proved.

THEOREM 2. The function $u \in M$ is a solution of (4) if and only if $u \in M$ satisfies the relation

$$u = P_M(u-\xi\Lambda(Tu-A(u)), \qquad \text{for all } v \in M, \tag{7}$$

where P_M is the projection of H into M and ξ is a positive constant.

PROOF. Suppose that $u \in M$ satisfies (4). Then by (5) this is equivalent to finding $u \in M$ such that

$$(\Lambda(Tu-A(u)), v-u) \geq 0, \qquad \text{for all } v \in M.$$

Now by invoking lemma 2 and for some constant $\xi>0$, the problem of finding $u \in M$ satisfying (4) is equivalent to finding $u \in M$ such that

$$u = P_M(u-\xi\Lambda(Tu-A(u)),$$

which is the required result.

Theorem 2 enables us to find the solution $u \in M$ satisfying (4) by the following iterative scheme.

$$u_{n+1} = P_M(u_n-\xi\Lambda(Tu_n-A(u_n)), \tag{8}$$

for some positive constant .

We now prove the strong convergence of the approximate solution to the exact solution.

THEOREM 3. Let u and u_{n+1} be solutions of (4) and (8) respectively. If the condition N holds, then

$$u_{n+1} \longrightarrow u \qquad \text{strongly in } H,$$

for $0<\xi<2(\rho-\gamma)/(\mu^2-\gamma^2)$ and $\gamma\xi<1$, where μ is the Lipschitz constant of T.

PROOF. By theorem 2, we know that $u \in M$ satisfying (4) is also a solution of (7) and conversely, Thus from (7) and (8), we obtain

$$\begin{aligned} \|u_{n+1}-u\| &= \|P_M(u_n-\xi\Lambda(Tu_n-A(u_n))-P_M(u-\xi\Lambda(Tu-A(u))\| \\ &\leq \|u_n-u-\xi\Lambda(Tu_n-Tu)+\xi\Lambda(A(u_n)-A(u))\|, \text{ see[7]}. \\ &\leq \|u_n-u-\xi\Lambda(Tu_n-Tu)\| + \xi\|A(u_n)-A(u)\|. \end{aligned}$$

Now following the technique of Noor[8], we immediately have

$$\|u_{n+1}-u\| \leq \theta\|u_n-u\|,$$

where $\theta = \sqrt{(1+\xi^2\mu^2-2\xi\rho)} + \gamma\xi < 1$ for $0<\xi<2(\rho-\gamma)/(\mu^2-\gamma^2)$ and $\xi\gamma<1$. Thus it follows from the above relation that u_{n+1} does converge strongly to u, the solution of (4), from the fixed point theorem. We also note that theorem 3 shows the existence of a unique solution of problem 1.

From theorem 3, we get a natural algorithm to compute the solution as follows;

i. $u_0 \in M$ is given.

ii. $u_{n+1} = P_M(u_n - \xi\Lambda(Tu_n - A(u_n)),$

where ξ is required to satisfy the condition $0<\xi<2(\rho-\gamma)/(\mu^2-\gamma^2)$ and $\gamma\xi< 1$.

REMARK 2.

A different class of variational inequalities known as quasi-variational inequalities has been considered in [2a]. Furthermore, the fixed point iterations suggested in [2a,p.79] and [1b] are only suitable for nonlinear quasi-variational inequalities. For difference and comparison of these two classes, see[1a].

REFERENCES

1a. C. Baiocchi and A. Capelo;Disequazioni variazionali e quasi variazionali, Applicazioni a problemi di frontiera libera, Vol.I&II, Bologna, 1978.

1b. L. Boccardo and I. Capuzzo Dolcetta;Esistenza e omogeneizzazione delle soluzioni di alcune disequazioni quasi variazionali nonlineari, to appear.

1. F.E. Browder;Nonlinear monotone operators and convex sets in Banach spaces, Bull. Amer. Math. Soc., 71(1965), 780-785.

2a. I. Capuzzo Docetta and U. Mosco;Implicit complementarity problems and quasi-variational inequalities in Variational Inequalities and Complementarity Problems (ed.) by R. Cottle, F. Giannessi and J. Lions, John Wiley & Sons, New York, (1980),75-87.

2. G. Duvaut and J. Lions; Inequations in Physics and Mehanics, Springer Verlag, Berlin, New York, 1976.

3. R. Glowinsli, J. Lions and R. Tremoliers;Numerical analysis of variational inequalities, North-Holland, 1981.

4. J. Lions and G. Stampacchia;Variational inequalities, Comm. Pure Appl. Math. 20(1967), 493-519.

5. E. Miersemann;Zur losungsverzweigung bei variationsungleichungen mit einer anwendung auf den knickstab mit bergenzter durchbiegung, Math. Nachr. 198(1981),7-15.

6. U. Mosco;An introduction to the approximate solution of variational inequalities in Constructive Aspects of Functional Analysis, Roma,(1973),499-685.

7. M. Aslam Noor;On Variational Inequalities, Ph.D. Thesis, Brunel Univ. U.K.1975.

8. M. Aslam Noor;Variational inequalities and approximation, Punjb. Univ. J. Math. 8(1975),24-40.

9. M. Aslam Noor;Strongly nonlinear problems in Hilbert space,BIMS,8(1980),3-8.

10. M. Aslam Noor and K. Inayat Noor;Minimum of nonlinear functionals on convex sets, Proc. Nat. Math. Conf. Tehran, (1977),80-89.MR80c ≠49008.

11. K. Inayat Noor and M.Aslam Noor;A generalization of the Lax-Milgram Lemma, Canad. Math. Bull. 23(1980),179-184. MR81h ≠47047.

12. R. Toscano and A. Maceri;On the elastic stability of beams under unilateral constraints; Meccanica, March(1980), 31-38.

13. M. Schechter,J. Shapiro and M. Snow;Solution of the nonlinear problem Au=N(u) in a Banach space,Trans. Amer. Math. Soc. 241(1978),69-78.

Received May 10, 1982

Mathematics Department,
King Saud University, Riyadh,
P.O.Box 2455, Saudi Arabia.

C.R. Math. Rep. Acad. Sci. Canada - Vol. IV, No. 4, August 1982 Août

RADAR RECEPTION AND NILPOTENT HARMONIC ANALYSIS III

Walter Schempp

Presented by P. Scherk, F.R.S.C.

Let $f \in L^2(\mathbf{R}^n)$ denote the pulse envelope of a modulated signal. Suppose that the signal has unit energy, i.e., that f admits L^2-norm $\|f\| = 1$. Then the radar autoambiguity function $H(f;.,.)$ with respect to f in the sense of Woodward is defined by the formula

$$H(f;x,y) = \int_{\mathbf{R}^n} f(t+{}^1/_2x)\bar{f}(t-{}^1/_2x)e^{2\pi i\langle y|t\rangle}dt$$

for all pairs $(x,y) \in \mathbf{R}^n \oplus \mathbf{R}^n$. This form of the radar autoambiguity function shows that $H(f;.,.)$ is obtained by correlating the signal of envelope f with its time-translated and Doppler-shifted version; that is, $H(f;.,.)$ is the correlation function on $\mathbf{R}^n \oplus \mathbf{R}^n$ in delay and Doppler. In theoretical optics, H is known as the indeterminacy or spread function. Observe that the double Fourier transform of $H(f;.,.)$ is the Wigner quasiprobability distribution function of a non-relativistic quantum-mechanical system with phase space $\mathbf{R}^n \oplus \mathbf{R}^n$ corresponding to the wavefunction $f \in L^2(\mathbf{R}^n)$. As we have seen in the first part [4] of this series of papers, the Wigner-Woodward relief (or radar ambiguity surface) admits symplectic symmetry. Indeed, let $f' \in L^2(\mathbf{R}^n)$ denote another envelope of L^2-norm $\|f'\| = 1$ and suppose that there exists for any pair of vectors $(x,y) \in \mathbf{R}^n \oplus \mathbf{R}^n$ a pair $(x',y') \in \mathbf{R}^n \oplus \mathbf{R}^n$ such that

$$H(f;x,y) = H(f';x',y')$$

holds. Then there are a unitary operator T of $L^2(\mathbf{R}^n)$ which is unique up to multiple by the scalar operator $\zeta.\mathrm{id}_{L^2(\mathbf{R}^n)}$, where $\zeta \in \mathbf{T}$, and a (unique) symplectic linear automorphism $\sigma \in Sp(n,\mathbf{R})$ such that the identities

$$(x,y) = \sigma(x',y'), \quad f = T(f')$$

hold. This solves the synthesis problem for radar autoambiguity functions with respect to signals of envelope $f \in L^2(\mathbf{R}^n)$ and represents the correct statement of Theorem 2 of [4] (see Theorem 3 of [1]). The signal transformation T has the following group-theoretical meaning. Consider the symplectic group $Sp(n,\mathbf{R})$ as a subgroup of the automorphism group of the real Heisenberg nilpotent group $\tilde{A}(\mathbf{R}^n)$ in the natural way. Then the automorphisms $\sigma \in Sp(n,\mathbf{R})$ leave the center of $\tilde{A}(\mathbf{R}^n)$ pointwise fixed. Define the action

$$\tilde{U}^{\sigma} = \tilde{U}^{\sigma^{-1}},$$

where $\tilde{U}$ denotes the Schrödinger representation of $\tilde{A}(\mathbf{R}^n)$. Then T is an intertwining operator of $\tilde{U}$ and $\tilde{U}^{\sigma}$, i.e., we have

$$T^{-1} \circ \tilde{U} \circ T = \tilde{U}^{\sigma}.$$

Since an application of the notion of Langrangian subspace (or, equivalently, of the notion of polarization) enables us to compute the unitary isomorphism T explicitly in terms of partial Fourier transforms, we are in a position to determine for a given radar ambiguity surface all the "admissible" pulse envelopes $f \in L^2(\mathbf{R}^n)$ in the associated autoambiguity function $H(f;.,.)$. This procedure shows again the crucial rôle played by symplectic geometry in the field of radar synthesis.

1. Polarizations

To begin with, we shall recall some basic facts of harmonic analysis on nilpotent Lie groups having discrete series. Let (V,B) denote a symplectic vector space of dimension $2n$, i.e., a vector space over the field $\mathbf{R}$ equipped with a nondegenerate alternating bilinear form B. Then the symplectic group $Sp(V,B)$ is formed by all linear automorphisms of V which preserve $B \in \Lambda^2 V$.

Let E denote an element $\neq 0$ of $\mathbf{R}$ and suppose that the $(2n+1)$-dimensional real vector space $\mathfrak{n} = V \oplus \mathbf{R}E$ is endowed with the structure of a Heisenberg algebra with center $\mathbf{R}E$ in the natural way. Then we have $[X,Y] = B(X,Y)E$ for all pairs $(X,Y) \in V \times V$. Let $\lambda \in \mathfrak{n}^*$ be the R-linear form such that $\langle E,\lambda\rangle = 1$ and denote by $B_\lambda \in \Lambda^2\mathfrak{n}$ the extension of B to $\mathfrak{n}$ defined by $B_\lambda(X,Y) = \langle [X,Y],\lambda\rangle$. Then a subalgebra $\mathfrak{l}$ of $\mathfrak{n}$ is called a polarization of $\mathfrak{n}$ associated with λ if $\mathfrak{l}$ forms a totally isotropic vector subspace of $\mathfrak{n}$ with respect to B_λ of maximal dimension n+1.

Let N denote the simply connected Heisenberg group of dimension 2n+1 associated with the Lie algebra $\mathfrak{n}$. Then the exponential mapping $\exp: \mathfrak{n} \longrightarrow N$ forms a diffeomorphism and Sp(V,B) acts on $\mathfrak{n}$ and also on N by automorphisms.

Define the function $\varepsilon : N \longrightarrow T$ by $\varepsilon(\exp X) = e^{2\pi i\langle X,\lambda\rangle}$ where $X \in \mathfrak{n}$ and introduce the induced representation

$$U_{\mathfrak{l}} = \operatorname{ind}_{\exp \mathfrak{l} \uparrow N} (\varepsilon \mid \exp \mathfrak{l}).$$

Then $U_{\mathfrak{l}}$ forms an irreducible unitary linear representation of N which acts on the complex Hilbert space $\mathcal{H}_{\mathfrak{l}}$. Notice that $(U_{\mathfrak{l}}, \mathcal{H}_{\mathfrak{l}})$ belongs to the discrete series of N.

Consider a pair $(\mathfrak{l}_1, \mathfrak{l}_2)$ of polarizations of $\mathfrak{n}$ associated with the same linear form $\lambda \in \mathfrak{n}^*$ we have defined above. According to the Stone-von Neumann-Segal theorem, the unitary linear representations $U_{\mathfrak{l}_1}$ and $U_{\mathfrak{l}_2}$ of N acting on the complex Hilbert spaces $\mathcal{H}_{\mathfrak{l}_1}$ and $\mathcal{H}_{\mathfrak{l}_2}$, respectively, are unitarily isomorphic. Let $\mu_{2,1} \neq 0$ denote a positive measure on $\exp \mathfrak{l}_2/\exp \mathfrak{l}_1 \cap \mathfrak{l}_2$ which is invariant under the action of $\exp \mathfrak{l}_2$ and standardized such that the linear integral operator (partial Fourier transform) given by

$$f \rightsquigarrow \left(N \ni x \rightsquigarrow \int_{\exp \mathfrak{l}_2/\exp \mathfrak{l}_1 \cap \mathfrak{l}_2} f(xl)\,\varepsilon(l)\,d\mu_{1,2}(l)\right)$$

may be extended to a unitary linear mapping $T_{\mathfrak{l}_2,\mathfrak{l}_1} : \mathcal{H}_{\mathfrak{l}_1} \rightarrow \mathcal{H}_{\mathfrak{l}_2}$. Then $T_{\mathfrak{l}_2,\mathfrak{l}_1}$ forms a unitary isomorphism of $U_{\mathfrak{l}_1}$ onto $U_{\mathfrak{l}_2}$ (cf. Lion [3]). In particular, for any polarization $\mathfrak{l}$ of $\mathfrak{n}$ associated with $\lambda \in \mathfrak{n}^*$ and any symplectic automorphism $\sigma \in Sp(V,B)$ the vector subspace $\mathfrak{l}_2 = \sigma(\mathfrak{l}_1)$ of $\mathfrak{n}$ forms a polarization of $\mathfrak{n}$ associated with λ. For any $f \in \mathcal{H}_{\mathfrak{l}_1}$ define $f^\sigma = f \circ \sigma^{-1} : N \ni x \rightsquigarrow f(\sigma^{-1}x)$. Then it is easy to see that $f \rightsquigarrow f^\sigma$ maps $\mathcal{H}_{\mathfrak{l}_1}$ onto $\mathcal{H}_{\mathfrak{l}_2}$ unitarily.

2. The Radar Synthesis Problem

Let us choose a canonical symplectic basis $(P_1,\ldots,P_n,Q_1,\ldots,Q_n)$ of the real vector space V. Then V can be identified with the phase space $R^n \oplus R^n$, Sp(V,B) can be realized as the matrix group Sp(n,R) and N can be realized as the real Heisenberg nilpotent group $\tilde{A}(R^n)$ in its basic presentation. Fix the polarization

$$\mathfrak{l} = \bigoplus_{1 \leq j \leq n} RQ_j \oplus RE$$

of $\mathfrak{n}$ associated with λ. Then U_λ can be realized as the Schrödinger representation $\tilde{U}$ of $\tilde{A}(R^n)$ (cf. [4]) and $\mathcal{H}_{\mathfrak{l}}$ as the complex Hilbert space $L^2(R^n)$. In view of Theorem 1 of [4] we obtain the following result.

<u>Theorem.</u> Let the envelopes $f \in L^2(R^n)$ and $f' \in L^2(R^n)$ be given such that $\|f\| = \|f'\| = 1$. Suppose that there exists for any pair of vectors $(x,y) \in R^n \oplus R^n$ a pair $(x',y') \in R^n \oplus R^n$ such that

$$H(f;x,y) = H(f';x',y')$$

holds. Then there exists a (unique) symplectic linear automorphism $\sigma \in Sp(n,R)$ satisfying

$$(x,y) = \sigma(x',y')$$

and

$$f = \zeta . T_{\mathfrak{l},\sigma\mathfrak{l}}(f'^{\sigma}),$$

where $\zeta \in T$ denotes a phase factor.

3. Examples

(i) Consider the symplectic automorphism

$$\sigma = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \in Sp(n,R)$$

with respect to a canonical symplectic basis of V (I_n = identity matrix). Then we have $T_{1,\sigma 1}(f'^{\sigma}) = \overline{\mathcal{F}}_{R^n} f'$ for all $f' \in L^2(R^n)$, i.e.,

$$H(f;x,y) = H(\zeta . \mathcal{F}_{R^n} f;-y,x)$$

holds for all pairs $(x,y) \in R^n \oplus R^n$ (see [4], Corollary 2 of Theorem 1). The right hand side expresses the radar autoambiguity function in terms of the spectrum $\mathcal{F}_{R^n} f$ of the envelope $f \in L^2(R^n)$ and an arbitrary phase factor $\zeta \in T$.

(ii) In the case n=1 we have obviously Sp(1,R) = SL(2,R). Suppose that the radar ambiguity surface H(f;R,R) with respect to the envelope $f \in L^2(R)$ is SO(2,R)-invariant. Then we may assume f' = W_m by the Corollary of Theorem 2 in [5] where W_m denotes the Hermite-Weber function of an arbitrary degree $m \geq 0$. We conclude by the Theorem that $f = \zeta . T_{1,\sigma 1}(f'^{\sigma}) = \zeta . W_m$ with $\zeta \in T$ for all $\sigma \in SO(2,R)$. Thus the radar ambiguity surface is invariant under rotations about the origin of the time-frequency plane and therefore secures simultaneously a high resolution in both range and range rate of a moving target if and only if the signal is a Hermite-Weber waveform whose envelope f is up to a phase factor an eigenfunction of the harmonic oscillator. Since the double Fourier transform of a radial function belonging to the space $L^2(R^2)$ is also radial, an analogous result holds in single-particle quantum mechanics for the rotational invariance of the Wigner quasiprobability distribution function corresponding to a wavefunction $f \in L^2(R)$. See Theorem 3 of [5]. Also see the papers by Klauder [2] which omits the details of the proofs and Wilcox [6] for a different approach to this example.

References

1. Hebsaker, H.-M., Schempp, W.: Radar detection, quantum mechanics, and nilpotent harmonic analysis. Meth. Verf. math. Phys. (in print)

2. Klauder, J.R.: The design of radar signals having both high range resolution and high velocity resolution. Bell System Tech. J. 39, 809-820 (1960)

3. Lion, G.: Extensions de représentations de groupes de Lie nilpotents et indices de Maslov. C. R. Acad. Sc. Paris 288, A 615 - A 618 (1979)

4. Schempp, W.: Radar reception and nilpotent harmonic analysis I. C. R. Math. Rep. Acad. Sci. Canada 4, 43 - 48 (1982)

5. Schempp, W.: Radar reception and nilpotent harmonic analysis II. C. R. Math. Rep. Acad. Sci. Canada 4 (1982)

6. Wilcox, C.H.: The synthesis problem for radar ambiguity functions. The University of Wisconsin, Mathematics Research Center. Madison, Wisconsin. Tech. Summary Report 157 (1960)

Lehrstuhl für Mathematik I
Universität Siegen
Hölderlinstraße 3
D-5900 Siegen, W.-Germany

Received June 7, 1982

C.R. Math. Rep. Acad. Sci. Canada - Vol. IV, No. 4, August 1982 Août

CHARACTERS OF MOTION GROUPS OVER GF(2)

J. S. Frame

Presented by G. de B. Robinson, F.R.S.C.

Abstract. Generic degree formulas for characters and lower level character values are found for the motion group over GF(2) that is a semi-direct product of an abelian translation group by an orthogonal group over GF(2).

1. **Introduction.** The semi-direct product of an elementary abelian "translation group" T_n^σ of order $N^2 = 2^{2n}$ by an orthogonal group $O_{2n}^\sigma(2) = G_n^\sigma$ over GF(2), (with σ the sign + or -, or the number 1 or -1), defines the motion group M_n^σ, such that $G_n^\sigma = M_n^\sigma/T_n^\sigma$. We choose the quadratic invariant forms $Q_n^\sigma(Z)$ for G_n^σ to be

$$Q_n^+(Z) = \sum_{i=1}^{n} z_{2i-1}z_{2i}, \qquad Q_n^-(Z) = z_1^2 + z_2^2 + Q_n^+(Z) \tag{1.1}$$

using row vectors Z. We denote by J_n the direct sum of n transposition matrices interchanging z_{2i-1} and z_{2i}. Then the matrices $\hat{M}$ of G_{n+1}^σ satisfy the relations

$$\hat{M}\, J_{n+1} \hat{M}^T = J_{n+1} \qquad \text{or} \qquad \hat{M}^T = J_{n+1}\, \hat{M}^{-1}\, J_{n+1} \tag{1.2}$$

We then observe that M_n^σ is isomorphic with the subgroup $\hat{M}_n^\sigma$ of index $(2N-\sigma)(N+\sigma)$ in G_{n+1}^σ, whose matrices stabilize the column vector $(0^{2n+1},1)^T$. These matrices have the factorization

$$\hat{M} = \hat{T}\,\hat{A}, \qquad \hat{T} = \begin{bmatrix} I_{2n} & C & 0 \\ 0 & 1 & 0 \\ C^T J_n & q & 1 \end{bmatrix}, \qquad \hat{A} = \begin{bmatrix} A & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1.3}$$

with $A \in G_n$, C a 2n x 1 column vector, and $q = Q_n^\sigma(C)$. Matrices of the isomorphic group M_n^σ are obtained by stripping off the last rows and columns of these matrices. The subgroup T_n^σ is normal.

Generic degree formulas are found for all the absolutely irreducible complex (AIC) characters of M_n^{σ}, derived from those found for the orthogonal groups [4]. We check that Theorem 3.2 of [3] for orthogonal groups also gives the character values of M_n^{σ} on the transposition class t (or C_t). Formulas for level 1 and 2 character values on all classes are also found. To obtain values for ten associated pairs of level 2 characters, the character matrix of $M_1^- \cong S_4$ is multiplied by two 5-vectors of basic functions to yield ten extended level 2 characters, each including a single AIC level 2 character that is positive on class t.

The central factor group of the Weyl group E_8, isomorphic with $G_4^+ = O_8^+(2)$ has M_3^+ as a subgroup of index $(16-1)(8+1) = 135$. This is a monomial group of order $2^6 8!$ containing $G_3^+ \cong S_8$. The character tables of M_3^+ and M_3^- display patterns that provided the insights for the following observations and theorems. [1]

2. Degrees and character values on principal classes. Each coset $T_n^{\sigma} g_i$ of T_n^{σ} in M_n^{σ} contains N^2 elements in one or more M_n^{σ}-classes $C_{i\lambda}$ whose class sizes $|C_{i\lambda}|$ are all multiples of the size $|C_{i1}|$ of a first listed "principal class" that contains all the conjugates in M_n^{σ} of g_i. If g_i has level ℓ, it has $L^2 = 2^{2\ell}$ times as many conjugates in M_n^{σ} as in G_n^{σ}. Each class of G_n^{σ} is contained in a principal class of M_n^{σ}. All characters of M_n^{σ} induced by G_n^{σ}-characters vanish on all non-principal classes of M_n^{σ}.

We distinguish three types of AIC-characters of M_n^{σ}. Those of type 1 are the characters of the factor group G_n^{σ}, whose values on all elements of the coset $T_n^{\sigma} g_i$ are those for g_i in G_n^{σ}. Each type 2 representation of M_n^{σ} is a monomial representation with a character χ^{2j} whose restriction to G_n^σ (given by its values on principal classes) is the G_n^σ-character induced by a character χ^{Dj} of its subgroup $D_n = G_n^+ \cap G_n^-$ of index $(N-\sigma)N/2$. Since $G_{n-1} \cong D_n/\langle\tau\rangle$, where τ is the central transposition in D_n, the degree of each pair of associated type 2 M_n^σ-characters is $(N-\sigma)N/2$ times the known degree of a corresponding G_{n-1}-character. The sum of squares of type 2 degrees is $|G_n^\sigma|(N-\sigma)N/2$.

The remaining AIC characters χ^{3j} of M_n^σ are of type 3. The restriction to G_n^σ of χ^{3j} (given by character values on principal classes) is the character χ^{Mj*} induced in G_n^σ by a character χ^{Mj} of its subgroup M_{n-1}^σ of index $(N-\sigma)(N/2+\sigma)$. Hence the sum of squares of type 3 degrees is $|G_n^\sigma|(N-\sigma)(N/2+\sigma)$, and the sum of squares of all AIC degrees is $|G_n^\sigma|(1+N^2-1) = |M_n^\sigma|$. All AIC characters are accounted for, and their values on principal classes are those of induced characters of G_n^σ. In particular, the value χ_t^j on the transposition class of each AIC character χ^j of M_n^σ is given by the formula of Theorem 3.2 in [3].

Theorem 2.1. The transposition class multiplier for an AIC character χ^j of M_n^σ of level ℓ is

$$\omega_t^j = N(N-\sigma)\chi_t^j/\chi_1^j = \pm(N-\sigma+s^\sigma)N/L, \quad L = 2^\ell, \tag{2.1}$$

where s^σ denotes the sum of the zeros of the degree polynomial.

As for G_n^σ-characters, we designate an M_n^σ-character by a degree symbol consisting of a code word for a monic degree polynomial $P_\ell^j(N)$, divided by an integer independent of n. The q^{th} letter in the code word is h,k,g, or i, according as $N-2^{q-1}$, $N+2^{q-1}$, both, or neither are factors of $P_\ell^j(N)$. A character of M_n^+ and its mate in M_n^- have their symbols interchanged by replacing h by k and k by h. Two associated characters of M_n^σ have the same degree symbol, except that a bar over the last letter indicates the one negative on class C_t. Labels and degree symbols for the t-positive M_n^+-characters of levels 0 and 1, and their degrees for M_3^+ and the mates in M_3^- are:

Label:	1	3_1	s	r	b	c	
M_n^+-degree:	i	hh/6	ig/3	hik/6	hN/2	hk/2	
M_3^+-degree:	1	7	20	14	28	35	(2.2)
M_3^--degree:	1	15	20	6	36	27	

The type 2 characters of level 1 are b and $\bar{b}$, while those of type 3 are c and $\bar{c}$. The character 1+b+c vanishes on non-principal classes of M_n^σ, since it is the M_n^σ-character induced by 1 in G_n^σ. The degree symbol for each G_{n-1}-character of level $\ell > 1$ becomes the degree symbol for the induced M_n^σ-character by prefixing h (for M_n^+) or k (for M_n^-), suffixing N, and dividing by $2^{2\ell-1}$. Similarly, the degree symbol for a character of M_{n-1}^+ (or M_{n-1}^-) becomes the symbol for a type 3 character of M_n^+ (or M_n^-) by replacing an initial i by hk (or kh) or an initial h by hg (or k by kg), and dividing by $2^{2\ell-1}$.

3. Character values for levels 1 and 2. We denote by $(\alpha+\sigma\beta)/2$ the character of G_n induced by the 1-character of G_n^σ, and by $\bar{\gamma} = I\gamma$ the class function (independent of n) that counts, as exponent of -2, the number of indecomposables over GF(2), in a decomposed matrix similar to a G_n-matrix of the given class, having cube roots of unity as eigenvalues in GF(4). Then the t-positive type 1 M_n^σ-characters of level 1 are

$$3_1 = (\alpha-3\sigma\beta+2\gamma)/6, \quad s = (\alpha-\gamma)/3 - 1, \quad r = (\alpha+3\sigma\beta+2\gamma)/6 - 1. \quad (3.1)$$

Restriction to M_n^σ of the level 1 characters 3_1 and r of G_{n+1}^σ yields respectively the reducible level 1 characters 3_1+b and r+1+c of M_n^σ, where b is of type 2, c of type 3, and 3_1,r and 1 are type 1 characters having equal values on classes C_{i1} and $C_{i\lambda}$. To evaluate b and c, we denote by $d\alpha$ (or $d\beta$, etc.) the class function for M_n^σ whose value on $C_{i\lambda}$ is the difference between the value of α (or β, etc.) on the class C_λ of G_{n+1}^σ containing $C_{i\lambda}$ and the value on the principal class C_{i1} of M_n^σ in the T_n^σ-coset with $C_{i\lambda}$.

<u>Theorem 3.1. The level 1 M_n^σ-characters b of type 2 and c of type 3 are expressible on all M_n^σ-classes by the formulas</u>

$$b = d\alpha/6 - \sigma d\beta/2, \qquad c+1 = d\alpha/6 + \sigma d\beta/2 \tag{3.2}$$

Setting $\alpha_2(g_i) = \alpha(g_i^2)$, etc., we next introduce four level 2 class functions of degree 0, derived from b and c:

$$\phi_b = (c+1)^{[1^2]} - (c+1)b + b^{[1^2]} = (d\beta)^2/2 - d\alpha_2/6 \tag{3.3}$$

$$\phi_c = (c+1)^{[2]} - (c+1)b + b^{[2]} - \alpha = (d\beta)^2/2 + d\alpha_2/6 - \alpha \tag{3.4}$$

$$f_b = d\beta_2^* - d\beta_2, \qquad f_c = d\beta_2^* - \beta \tag{3.5}$$

where β_2^* is either β_2 or 0 according as the class label has an even or odd numbers of factors t.

<u>Theorem 3.2. Generic formulas for the five t-positive type 2 AIC M_n^+-characters of level 2, and their degrees for M_3^+ are, for $\sigma = 1$:</u>

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 3 & -1 & 0 & 1 & -1 \\ 2 & 2 & -1 & 0 & 0 \\ 3 & -1 & 0 & -1 & 1 \\ 1 & 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} b\alpha/4! \\ \phi_b/8 \\ b\gamma/3 \\ \sigma b\beta/4 \\ \sigma f_b/4 \end{bmatrix} = \begin{bmatrix} hkkN/48 \\ hhkN/16 + b \\ hgN/24 \\ hkhN/16 + b \\ hhhN/48 \end{bmatrix}, \quad \begin{bmatrix} 140 \\ 252 \; + 28 \\ 140 \\ 140 \; + 28 \\ 28 \end{bmatrix} \tag{3.6}$$

Corresponding M_n^--characters are obtained with $\sigma = -1$, which interchanges h's and k's. The M_3^- degrees are 36,180,180, 324, and 180.

Theorem 3.3. Generic formulas for the five t-positive type 3 AIC M_n^+-characters of level 2, and their degrees for M_3^+ are, for $\sigma = 1$:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 3 & -1 & 0 & 1 & -1 \\ 2 & 2 & -1 & 0 & 0 \\ 3 & -1 & 0 & -1 & 1 \\ 1 & 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} c\alpha/4! \\ \phi_c/8 \\ c\gamma/3 \\ \sigma c\beta/4 \\ \sigma f_c/4 \end{bmatrix} = \begin{bmatrix} hgik/48 + c \\ hgk/16 + c \\ hkg/24 + c \\ hgN/16 \\ hgh/48 \end{bmatrix}, \quad \begin{bmatrix} 140 + 35 \\ 315 + 35 \\ 140 + 35 \\ 210 \\ 35 \end{bmatrix} \qquad (3.7)$$

Corresponding M_n^--characters are obtained with $\sigma = -1$, which interchanges h's and k's. The M_3^- degrees are 0, 135, 108, 270, and 135.

The formulas of Theorems 3.2 and 3.3, involving the character matrix of $M_1^- = S_4$, resemble those in [2] for the level 2 orthogonal group characters. Here the known level 1 characters b or c must be extracted from some of the level 2 extended characters to obtain the level 2 AIC-characters of types 2 and 3.

References

1. J.S.Frame, The characters of the Weyl group E_8. Computational Problems in Abstract Algebra, edited by John Leech. Pergamon Press, 1970, pp. 111-130.

2. J.S.Frame, Character formulas for the orthogonal groups over GF(2) II. Mathematical Reports, vol. 1, no. 3 (1979), 153-156.

3. J.S.Frame, Degree formulas for the orthogonal groups over GF(2). Mathematical Reports, vol. 2, no. 4 (1980), 199-202.

4. J.S.Frame, Degree polynomials for the orthogonal groups over GF(2). Mathematical Reports, vol. 2, no. 5 (1980), 253-258.

Michigan State University
East Lansing, Michigan

Received June 14, 1982

C.R. Math. Rep. Acad. Sci. Canada - Vol. IV, No. 4, August 1982 Août

A NEW THEORY OF LINEAR TREND WHICH RESOLVES THE GMR DEBATE AND SOLVES THE ERRORS-IN-VARIABLES PROBLEM

Jon Schnute

Presented by M. Shinbrot, F.R.S.C.

ABSTRACT. This paper summarizes a new theory of linear trend for bivariate data based on general properties of the error, including scale-invariance, symmetry, and normality. The theory makes possible an objective resolution of a debate over geometric mean regression (GMR), and it provides a method for solving the errors-in-variables problem without recourse to information outside the sample data.

INTRODUCTION. The geometric mean regression (GMR) defines a line through the centroid of bivariate data with a slope equal in magnitude to the ratio of standard deviations for y and x. The name derives from the fact that this slope is the geometric mean of slopes for the familiar y-on-x and x-on-y regressions. Ricker (1973), following Teissier (1948) and others, espoused the use of GMR in various aspects of fishery research for several reasons, in particular because the slope estimate gives no preference to x or y. He soon encountered stiff opposition to this point of view. Critics pointed out that, among other things, Ricker seemed to ignore a well-known problem in estimating the slope of a line when the variables are measured with error: consistent likelihood estimates are impossible without extra information beyond the data.

This paper examines the concept of a linear trend through bivariate data by considering the simplest possible definitions of error (departure from trend). GMR finds a natural place in this new theory as the answer to a reasonable basic question. Perhaps even more significantly, the theory gives an estimator for the slope of a line in the errors-in-variables context which at least meets the criterion of statistical consistency without requiring information beyond the data.

DEFINITIONS. Let Z(X,Y,a) be a scalar random variable determined by the random pair (X,Y) and the real number a. Then Z is called a _linear error_ if (i) Z depends linearly on X and Y and (ii) Z vanishes if and only if (X,Y) lies on the line of slope a through the mean of (X,Y). Z is _continuous_ if (i) Z depends continuously on a for a>0 and (ii) Z(-X,Y,-a) = Z(X,Y,a). Two linear errors are _equivalent_ if they differ only by a non-zero multiplicative constant independent of (X,Y) and a. A linear error Z is _scale-invariant_ if, for any pair (p,q) of non-zero constants, the transformations (X,Y) → (pX,qY) and a → qa/p carry Z to an equivalent error. Z is _symmetric_ if the transformations (X,Y) → (Y,X) and a → 1/a carry Z to an equivalent error.

Given a random pair (X,Y) and a linear error Z, let V(a) represent the variance Var[Z(X,Y,a)], and let V' be the infimum of V(a) over real a. Then, if there exists a unique a' such that V(a')=V', a' is called the _variance-optimal slope_ for (X,Y) and Z.

A random pair (X,Y) is said to have a _trend line_ of slope a" through the mean E[(X,Y)] if there exists a unique slope a'', called the trend slope, for which $Z(X,Y,a'')$ is normal for any linear error Z. The problem of finding a consistent estimator for a'' from a finite sample drawn from (X,Y) is called the trend estimation problem.

The random pair (X,Y) is an N-component binormal mixture if

$$(X,Y) = \Sigma W_i(U_i,V_i),$$

where (U_i,V_i) for $i=1,\ldots,N$ are independent identically distributed binormal random pairs independent of the scalar random variables W_j for $j=1,\ldots,N$. The component means $E[(U_i,V_i)]$ are presumed distinct, and the random weights W_i are assumed to be constrained by $\Sigma W_i = 1$ and $\Sigma W_i{}^2 = C$, where C is a constant with $C>1/N$. The support of the random vector $(W_1,\ldots,W_N)$ is assumed to contain at least N independent points in R^N. The mixture is uncoupled if $(W_1,\ldots,W_N)$ can take only values with one non-zero component; otherwise it is coupled. It is fully coupled if, for each i, W_i vanishes with zero probability. It has small errors if $Var(U_1)$ and $Var(V_1)$ are small compared to $Var(X)$ and $Var(Y)$, respectively.

The random pair (X,Y) is a linear scatter if $(X,Y) = (S,T) + (U,V)$, where (i) (U,V) is binormal with mean $(0,0)$, (ii) (S,T) is a random pair with a support which is confined to a line and contains at least two points, and (iii) (S,T) and (U,V) are independent. The line determined by the support of (S,T) is called the scatter line. (X,Y) has small errors if $Var(U)$ and $Var(V)$ are small compared to $Var(X)$ and $Var(Y)$, respectively. The problem of finding a consistent estimator for the slope of the scatter line from a finite sample is the <u>errors-in-variables problem</u>.

THEOREM 1. There is a natural map between real numbers k and equivalence classes of continuous scale-invariant linear errors Z, where Z(k) is also symmetric if and only if k=0. Furthermore, V'=0 for arbitrary (X,Y) unless $-1 \leq k \leq 1$. If Cov(X,Y) is nonzero, then the variance-optimal slope a'(k) exits for (X,Y) and scale-invariant Z(k), and the slopes a'(-1), a'(0), and a'(1) correspond naturally to the x-on-y, geometric mean, and y-on-x regression slopes, respectively.

CORROLLARY. The GMR slope is the sample moment estimate of the variance-optimal slope for scale-invariant symmetric linear error.

THEOREM 2. An N-component binormal mixture has a trend line if and only if the component means are colinear. In this case, the line of means is the trend line.

THEOREM 3. If (X,Y) is a linear scatter, then either (X,Y) is binormal or the scatter line is a trend line for (X,Y).

THEOREM 4. In Theorems 2 and 3, if (X,Y) has a trend line and also has small errors, then, to first order terms in small variance ratios,

$$a'(k) \cong a'' (1 + \varepsilon_1 + \varepsilon_2 k),$$

where a'(k) with $-1 \le k \le 1$ and a" are the variance-optimal and trend slopes, respectively, and ε_1 and ε_2 are small quantities independent of k.

THEOREM 5. The trend estimation problem is solvable and, in particular, so is the errors-in-variables problem for a non-binormal linear scatter.

DISCUSSION. Theorem 1 places GMR in a natural context between x-on-y and y-on-x regressions. The constant k relates to the dimensionality of the error. For the cases k = -1, 0, and 1, the units of Z are, essentially, [X], $[XY]^{1/2}$, and [Y], respectively. Theorems 2 and 3 give examples of random pairs with trend lines, and Theorem 4 gives a context in which the variance-optimal slope (for any k) is approximately equal to the trend slope.

Theorem 5 is based on an estimator that measures normality of error, the trend slope estimate being the one with most apparently normal error Z(X,Y,a). Typically, this estimator involves skewness and kurtosis measurements, that is, the third and fourth order moments of the random pair (X,Y). The textbook likelihood slope estimate for the errors-in-variables problem involves only moments of order 2, and it isn't difficult to show that these moments alone do not carry enough information to estimate trend slope. On the other hand, a likelihood approach to the linear scatter problem would also give estimates of the realizations (s,t) and (u,v) which produce each observed data point (x,y). The theory here focuses only on the trend line as a whole, without concern for identifying a "true" point on the line with each observed data point. The fully coupled N-component colinear mixture distribution affords an example for which, in fact, there is no one point on the line corresponding to each data point.

A detailed account of the above work, including proofs of theorems and derivations of estimators and their variances, appears in Schnute (1982).

ACKNOWLEDGEMENTS. I undertook this project thanks to the advice, encouragement, and prodding of Dr. W. E. Ricker, F.R.S.C. I am also much in debt to Prof. Bill Reed at the University of Victoria, B.C., for a sensitive reading of my work, and, in particular, for suggesting a version of Theorem 3 as a bonus to the methods of proving Theorem 2. Finally, I am grateful to Prof. M. Shinbrot, F.R.S.C., for his active review of this work and his helpful suggestions on the method of presentation.

Ricker, W. E. Linear regressions in fishery research. J. Fish. Res. Board Can. 30 (1973): 409-434.

Schnute, J. A new theory of linear trend based on the GMR debate. Submitted to J. Amer. Stat. Assoc. (1982).

Teissier, G. La relation d'allometric: sa signification statistique et biologique. Biometrics 4 (1948): 14-48.

Jon Schnute
Pacific Biological Station
Nanaimo, B.C. V9R 5K6

Received June 28, 1982

C.R. Math. Rep. Acad. Sci. Canada - Vol. IV, No. 4, August 1982 Août

ON BRAUER CHARACTERS

B.M. Puttaswamaiah

Presented by G.de B. Robinson, F.R.S.C.

There is no known single method of derivation of orthogonality relations for irreducible characters which encompasses both the ordinary and Brauer characters. All the known derivations depend on a knowledge of the structure theory of finite dimensional algebras over complete local fields. For example, see [4]. By generalizing the method of ordinary characters given in [1], we obtain the orthogonality relations for Brauer characters. The method avoids the use of structure theory of finite dimensional algebras over complete fields. Although the results are not new, the method is believed to be new. As an application, a result of Burnside-Brauer is generalized.

Let G be a group of finite order g, p be a fixed positive rational prime, G^o be the set of all p-regular elements of G, $C_1 = \{1\}$, $C_2,\ldots,C_n$ be the p-regular conjugate classes of G, $C_{n+1}, C_{n+2},\ldots,C_s$ be the remaining conjugate classes of G, g_i the number of elements in C_i, $C\ell(G^o)$ be the algebra of all class functions from G^o into the complex field K, $\zeta^1, \zeta^2, \ldots, \zeta^s$ be the irreducible ordinary characters of G, $\varphi^1, \varphi^2, \ldots, \varphi^n$ be the (absolutely) irreducible Brauer (modular) characters of G at the fixed prime p, Φ be the regular character of G, and KG be the group algebra of G. For any character θ^i, θ^{i*} denotes the character given by $\theta^{i*}(x) = \overline{\theta^i(x)}$, for all x in G, where the bar denotes the complex conjugate. There are several equivalent definitions of Brauer characters [For example see 2,3]. We assume that ζ^1 and φ^1 are the irreducible trivial characters of G. These notations will be kept fixed throughout this paper.

By restricting the domain (if necessary), we may assume that an irreducible ordinary character always lies in $C\ell(G^o)$. The Brauer characters φ^i are linearly independent [2], and so they form a basis of $C\ell(G^o)$ over K. For any prime p, the restriction of ζ^i to G^o is a Brauer character of G[2, p. 265]. Therefore there are uniquely determined non-negative integers d_{ij} such that

(1) $$\zeta^i = \sum_{j=1}^{n} d_{ij}\,\varphi^j; \quad i = 1,2,\ldots, s.$$

The integers d_{ij} are the decomposition numbers of G at p and the s by n matrix $D = [d_{ij}]$ is the decomposition matrix of G at p. Since the irreducible Brauer characters are linearly independent, the matrix D has rank n. By renumbering (if necessary) except ζ^1, we may assume that $\zeta^1,\zeta^2,\ldots,\zeta^n$ considered as members of $C\ell(G^o)$ are linearly independent. Note that $d_{11} = 1$ and $d_{1j} = 0$ for all $j > 1$.

Using the decomposition numbers, we define

(2) $$\eta^i = \sum_{j=1}^{s} d_{ji}\,\zeta^j; \quad i = 1,2,\ldots,n.$$

Then η^i is an ordinary character. The characters η^i considered as members of $C\ell(G^o)$ are called the projective indecomposable characters of G at the prime p. The Brauer character φ^i is a constituent of η^i. Indeed by (1) and (2), we have

$$\eta^i = \sum_{j=1}^{s} d_{ji}\,\zeta^j = \sum_{k=1}^{n}\Big(\sum_{j=1}^{s} d_{ji}d_{jk}\Big)\varphi^k.$$

The coefficient of φ^i is $\sum_{j=1}^{s} d_{ji}d_{ji} = \sum_{j=1}^{s} d_{ji}^2 \neq 0$, since the rank of D is n.

<u>Lemma 1</u>. With the above notation

(3) $$\sum_{i=1}^{s} \zeta^i(x)\,\zeta^i(y) = \sum_{i=1}^{n} \varphi^i(x)\eta^i(y); \quad x \in G^o$$

Proof:

$$\sum_{i=1}^{s} \zeta^i(x)\, \zeta^i(y) = \sum_{i=1}^{s} \sum_{j=1}^{n} d_{ij}\, \varphi^j(x)\, \zeta^i(y)$$
$$= \sum_{j=1}^{n} \varphi^j(x) \sum_{i=1}^{s} d_{ij}\, \zeta^i(y)$$
$$= \sum_{j=1}^{n} \varphi^j(x)\, \eta^j(y)\ .$$

Corollary 1. With the above notation $\eta^j(y) = 0$ for $y \in G\setminus G^o$.

Proof: This is a consequence of the orthogonality relations for irreducible ordinary characters, (3) and the linear independence of irreducible Brauer characters. Also see [2].

Corollary 2. With the above notation

(4) $$\sum_{x \in G^o} \eta^i(x) = g\, \delta_{1,i}$$

where $\delta_{i,j}$ is the Kronecker delta.

Proof: By Corollary 1, $\sum_{x \in G^o} \eta^i(x) = \sum_{x \in G} \eta^i(x)$ and so

$$\sum_{x \in G^o} \eta^i(x) = \sum_{x \in G} \eta^i(x) = \sum_{j=1}^{s} d_{ji} \sum_{x \in G} \zeta^j(x) = g\, d_{1,i} = g\, \delta_{1,i}\ .$$

Lemma 2. If Φ is the regular (complex) character of G, then

(5) $$\Phi = \sum_{i=1}^{n} z_i\, \eta^i$$

where $z_i = \varphi^i(1)$ is the degree of φ^i for $i = 1,2,\ldots,n$.

Proof: By (1), $\zeta^i(1) = \sum_{j=1}^{n} d_{ij}\, z_j$ and so

$$\sum_{i=1}^{n} z_i\, \eta^i = \sum_{j=1}^{s} \Big(\sum_{i=1}^{n} d_{ji}\, z_i\Big) \zeta^j = \sum_{j=1}^{s} \zeta^j(1)\, \zeta^j = \Phi.$$

Lemma 3. The projective indecomposable characters form a basis of $C\ell(G^o)$.

Proof: Let $\sum_{i=1}^{n} a_i\, \eta^i = 0$ where $a_i \in K$. By Corollary 1 of Lemma 1 and equation (2), we have

$$0 = \sum_{x\in G^o} \sum_{i=1}^{n} a_i \eta^i(x) \overline{\zeta^j(x)} = \sum_{i=1}^{n} a_i \sum_{k=1}^{s} d_{ki} \sum_{x\in G} \zeta^k(x) \overline{\zeta^j(x)}$$

$$= \sum_{i=1}^{n} \sum_{k=1}^{s} a_i d_{ki}\, g\, \delta_{k,j}$$

$$= g \sum_{i=1}^{n} a_i d_{ji} .$$

We can consider $\sum_{i=1}^{n} a_i d_{ji} = 0$ for $j = 1,2,\ldots,n$ as a system of n equations in n unknowns $a_1,a_2,\ldots,a_n$. By the remark made following the definition of d_{ij}, the coefficient matrix of the system is non-singular. Hence $a_1 = a_2 = \ldots = a_n = 0$. Since $\eta^1,\eta^2,\ldots,\eta^n$ are linearly independent, they form a basis of $C\ell(G^o)$ over K.

<u>Lemma 4</u>. For any i,j, $\eta^i\varphi^j$ is a linear combination of $\eta^1,\eta^2,\ldots,\eta^n$ with non-negative rational integer coefficients.

Proof: By (5), $\eta^i\varphi^j$ is a constituent of $\Phi\varphi^j$. But $\Phi\varphi^j = z_j\Phi = z_j(\sum_{j=1}^{n} z_i \eta^i)$ and hence $\eta^i\varphi^j$ is a linear combination of $\eta^1,\eta^2,\ldots,\eta^n$ with non-negative rational integer coefficients by Krull-Schmidt theorem [4].

<u>Lemma 5</u>. If η^1 is a constituent of $\eta^i\varphi^j$, then $i = j^*$.

Proof: By Lemma 4, there are unique non-negative rational integers a_{ijk} such that

$$\eta^i\varphi^j = \sum_{k=1}^{n} a_{ijk}\, \eta^k$$

where $a_{ij1} \neq 0$. By (5),

$$\Phi\varphi^j = \sum_{i=1}^{n} z_i\, \eta^i\varphi^j = \sum_{k=1}^{n} (\sum_{i=1}^{n} z_i\, a_{ijk})\eta^k,$$

so that

$$\sum_{x\in G^o} \Phi(x^{-1})\, \varphi^j(x^{-1}) = \sum_{k=1}^{n} (\sum_{i=1}^{n} z_i\, a_{ijk}) g\, \delta_{1,k} .$$

Hence $z_{j*} = \sum_{i=1}^{n} z_i\, a_{ij1}$. By (3) and (4), $a_{j*j1} \neq 0$ and so $j^* = i$ and $a_{ij1} = 0$ for $i \neq j^*$.

<u>Corollary</u>. For each i, η^1 is a constituent of $\eta^i\varphi^{i*}$ with multiplicity 1. [Compare this result with the result in [1] on page 137].

THEOREM (Orthogonality Relations). If $\eta_k^i = \eta^i(x)$ and $\varphi_k^i = \varphi^i(x)$ with $x \in C_k$, then

(6) $$\sum_{i=1}^{n} g_i \overline{\eta_i^k} \varphi_i^\ell = g\,\delta_{k,\ell} \quad \text{and} \quad (7) \quad \sum_{i=1}^{n} \overline{\eta_k^i} \varphi_\ell^i = g\,\delta_{k,\ell}/g_k,$$

for all k,ℓ.

Proof: First assume that $k \neq \ell$. Then we assert that

(8) $$\sum_{x \in G^o} \overline{\eta^k(x)}\, \varphi^\ell(x) = 0 .$$

Indeed, write $\overline{\eta^k(x)}\, \varphi^\ell(x) = \sum_{t=1}^{n} a_{k\ell t}\, \eta^t(x)$ with $a_{k\ell t}$ in K and $a_{k\ell 1} = 0$ for $k \neq \ell$. Summing over G^o, we get

$$\sum_{x \in G^o} \overline{\eta^k(x)}\, \varphi^\ell(x) = \sum_{t=2}^{n} a_{k\ell t} \sum_{x \in G^o} \eta^t(x) = 0,$$

so that the result (6) follows in this case.

Next by Lemma 2, $\Phi = \sum_{k=1}^{n} z_k \eta^k$ is the regular character of G, and so $\Phi(1) = g$ and $\Phi(x) = 0$ for $x \neq 1$. Then

$$\sum_{x \in G^o} \Phi(x)\, \varphi^\ell(x^{-1}) = \sum_{k=1}^{n} z_k \sum_{x \in G^o} \eta^k(x)\, \varphi^\ell(x^{-1})$$

so that by (8), we get

$$g\, z_\ell = \sum_{k \neq \ell} z_k \sum_{x \in G^o} \eta^k(x)\, \varphi^\ell(x^{-1}) + z_\ell \sum_{x \in G^o} \eta^\ell(x)\, \varphi^\ell(x^{-1})$$

$$= 0 + z_\ell \sum_{x \in G^o} \eta^\ell(x)\, \varphi^\ell(x^{-1}).$$

Hence

$$\sum_{x \in G^o} \eta^\ell(x) \varphi(x^{-1}) = g$$

which proves (6).

The relation (7) is a consequence of (6).

Since the irreducible Brauer characters are linearly independent, there are unique integers c_{ij} such that

(9) $$\eta^i = \sum_{j=1}^{n} c_{ij}\, \varphi^j; \quad i = 1,2,\dots,n .$$

The integers c_{ij} are the Cartan invariants of G at the prime p and $C = [c_{ij}]$ is the Cartan matrix of G at p. Then (1), (9) and linear independence of Brauer characters yield $C = D'D$ where the dash denotes the transpose. Clearly $c_{ij} = c_{ji}$ for all i,j. Moreover, C is non-singular. If $C^{-1} = [r_{ij}]$, then (6) and (7) imply that

$$\sum_{k=1}^{n} g_k \eta_k^i \overline{\eta_k^j} = c_{ij}\, g$$

and

$$\sum_{k=1}^{n} g_k \varphi_k^i \overline{\varphi_k^j} = g\, r_{ij}$$

for all $i,j = 1,2,\ldots,n$.

An application. In [5], Robinson raised the question of sharpening the result of Burnside on the power of a faithful irreducible ordinary character. Brauer gave a refinement [2, p. 49] of this result. The following result further improves this refinement.

Theorem . Let θ be an irreducible ordinary character of G, which takes on r distinct values. Then every irreducible character ψ of G such that $\operatorname{Ker}\theta \leq \operatorname{Ker}\psi$ is a constituent of θ^j for $0 \leq j \leq r$.

The proof of this result is similar to the one given in [2, p. 49].

REFERENCES

1. I.N. Herstein, Noncommutative Rings, The Carus Mathematical Monograph (1968).

2. I.M. Isaacs, Character Theory of Finite Groups, Academic Press, New York (1976).

3. B.M. Puttaswamaiah, Determination of Brauer Characters, Canadian Journal of Mathematics, 26(1974), 746-752.

4. B.M. Puttaswamaiah and John D. Dixon, Modular Representations of Finite Groups, Academic Press, New York (1977).

5. G. De B. Robinson, Geometry of group representations, Nagoya Mathematical Journal, Vol. 27, (1977).

Carleton University
Ottawa, Canada

Received June 30, 1982

1. J. Cizek	Dept. of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3Gl
2. B.R. Ebanks	Dept. of Mathematics, Texas Technical University, Lubbock, Texas 79409, U.S.A.
3. J.S. Frame	Dept. of Mathematics, Michigan State University, East Lansing, Michigan 48824, U.S.A.
4. P.L. Kannappan	Dept. of Pure Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3Gl
5. M.A. Noor	Dept. of Mathematics, King Saud University, Riyadh, P.O. Box 2455, Saudi Arabia
6. B.M. Puttaswamaiah	Dept. of Mathematics and Statistics, Carleton University, Ottawa, Ontario, Canada KlS 5B6
7. W. Schempp	Lehrstuhl für Math. I, Universität Siegen, Hölderlinstrasze 3, D-5900 Siegen, West Germany
8. J. Schnute	Pacific Biological Station, Nanaimo, British Columbia, Canada V9R 5K6
9. L. Székelyhidi	Dept. of Mathematics, University of Debrecen, H-4010, Debrecen Pf. 12, Hungary
10. E.R. Vrscay	Dept. of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3Gl