JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 57, 667–675 (1977)

The Common Causality Structure of Multilinear Maps and Their Multipower Forms*

WILLIAM A. PORTER

Computer, Information and Control Engineering, The University of Michigan, Ann Arbor, Michigan 48104

Submitted by C. L. Dolph

1. INTRODUCTION

This paper deals with the causality structure of multilinear, multipower operators on a Hilbert resolution space. One objective of this study is to clear up certain unanswered questions raised in [1]. In particular we present a counterexample consisting of a compact set $K \subset L_2(0, 1)$ such that no memoryless multipower operators of order $n \geqslant 2$ can be defined on K. This lays to rest the tempting conjecture that a Weierstrass-type approximation result holds between the finite memoryless polynomic functions and the memoryless continuous functions.

In order to present the counterexample it is necessary to prove that certain causality properties of a multipower operator hold if and only if the symmetric multilinear generator of the multipower operator also has these properties in each linear argument. These results are important in their own right.

Because this study is motivated by and supportive of [1], we shall adopt the notation, definitions, and conventions of this reference. In brief, H is a Hilbert space and $W: H^n \to H$ is a *multilinear* operator if $W[x_1, x_2, ..., x_n]$ is linear in each argument. W is symmetric if it is invariant under all possible permutations of arguments, for instance, $W[x_1, x_2, ...] = W[x_2, x_1, ...]$ all $x_1, x_2 \in H$. The multilinear function W generates a *multipower* function, $\hat{W}: H \to H$, by the formula $\hat{W}(x) = W[x, x, ..., x]$. If W generates $\hat{W}$ then there is a symmetric $\overline{W}$ determined by W which also generates $\hat{W}$ (see [1, 2, 4]). For this reason, we focus exclusively on multipower functions and their symmetric multilinear generators.

A family $\{P^t : t \in \nu\}$ of orthoprojectors on H is said to be a resolution of the identity if ν is a linearly ordered set with maximal and minimal elements t_∞, t_0, respectively, and if (1) $P^t(H) \supseteq P^l(H)$ whenever $t \geqslant l$, and (2) $P^{t_0}(H) = \{0\}$,

* Sponsored in part by the U. S. Air Force Office of Scientific Research, Grant No. 732427, and No. 77–0352.

ISSN 0022–247X

$P^{t_\infty}(H) = H$ hold. The Hilbert space H, equipped with the resolution $\{P^t\}$ is called a *Hilbert resolution space*. For later use we note that $L_2(0, \tau)$ equipped with the projections

$$(P^t x)(\beta) = x(\beta), \quad \beta \leqslant t, \qquad t, \beta \in [0, \tau],$$
$$= 0, \quad t < \beta,$$

is a Hilbert resolution space.

A function $f\colon H \to H$ is said to be *causal* (anticausal) if $P^t f = P^t f P^t$ all $t \in \nu$ $((I - P^t) f (I - P^t)$ all $t \in \nu)$, respectively. If f is both causal and anticausal it is called *memoryless*. The function is *prestrictly causal* if it is causal and moreover there exist finite sets $\{t_0 < t_i < \cdots < t_{N+1} = t_\infty\} \subset \nu$ and $\{\Delta_i = P^{i+1} - P^i : i = 0,..., N\}$ such that $f = \sum_{i=0}^N \Delta_i f P^{i-1}$. Here we adopt the abbreviation $P^i = P^{t_i}$. The function f is *strictly causal* if it is in the closure (uniform or strong) of the prestrictly causal class. For examples of functions with these various properties the reader is referred to [1, 5]. We note that on $L_2(0, \tau)$, indefinite integration is strictly causal (and causal) and scalar multiplication is memoryless (and causal, anticausal), while time translation is prestrictly causal (and strictly causal, causal).

If W is the multipower operator induced by multilinear W then it is easily shown that the causality properties of W determine those of $\hat{W}$. In particular. if W is causal (anticausal, memoryless, prestrictly causal, strictly causal) in each argument then $\hat{W}$ is also of the same type. If W is symmetric then it must have the same causality property in all arguments. In the following we show that these sufficient conditions are also necessary.

2. Some Algebraic Relations

Let W denote a symmetric multilinear generator of the multipower function $\hat{W}$. The relations between W and $\hat{W}$ exhibit most of the algebraic structure of the power functions on R. These similarities can be used to our advantage in simplifying our discussion.

The main property we shall need here is related to the identities

$$2! x_1 x_2 = (x_1 + x_2)^2 - (x_1)^2 - (x_2)^2,$$

$$3! x_1 x_2 x_3 = \left(\sum_{i=1}^3 x_i\right)^3 - \sum_{j=1}^3 \left(\sum_{i \neq j} x_i\right)^3 + \sum_{j=1}^3 (x_j)^3,$$

$$4! x_1 x_2 x_3 x_4 = \left(\sum_{i=1}^4 x_i\right)^4 - \sum_{j=1}^4 \left(\sum_{i \neq j} x_i\right)^4 + \sum_{j,k=1}^4 \left(\sum_{i \neq j,k} x_i\right)^4 - \sum_{j=1}^4 (x_j)^4,$$

which hold on the scalar field. These identities, which can be verified by direct inspection, have an apparent pattern. Indeed, for arbitrary $n > 0$

$$n!x_1 ,..., x_n = \left(\sum_{i=1}^{n} x_i\right)^n - \sum_{j=1}^{n} \left(\sum_{i\neq j} x_i\right)^n + \cdots + (-1)^{n-1} \sum_{j=1}^{n} (x_j)^n. \qquad (1)$$

We note that the map $W[x_1 ,..., x_n] = x_1 x_2 ,..., x_n$ is n-linear, symmetric and that $\hat{W}(x) = W[x,..., x] = (x)^n$. Equation (1) indicates that on the scalars multipower $\hat{W}$ can be used to compute multilinear W.

In an earlier study Schetzen [6] elevated Eq. (1) to operator form. Schetzen's result is embodied in our first lemma.

LEMMA 1. *Let W be the symmetric n-linear generator of n-power $\hat{W}$ on H. Then W can be computed using $\hat{W}$ by the formula*

$$n!W[x_1 ,..., x_n] = \hat{W}\left(\sum_{i=1}^{n} x_i\right) - \sum_{j=1}^{n} \hat{W}\left(\sum_{i\neq j} x_i\right) + \cdots + (-1)^{n-1} \sum_{j=1}^{n} \hat{W}(x_j). \qquad (2)$$

As with the power functions on the scalar field Lemma 1 can be verified by direct inspection.

3. CAUSALITY PROPERTIES

We now assume $\{H, P^t\}$ is a Hilbert resolution space and investigate the relative causality structure of multipower functions and their symmetric generators.

LEMMA 2. *Let symmetric n-linear W generate n-power $\hat{W}$. If $\hat{W}$ is causal, i.e., $P^t\hat{W} = P^t\hat{W}P^t$, $t \in \nu$, then $W[x_1 ,..., x_n]$ is causal in every argument.*

Proof. Using Lemma 1 we have

$$\begin{aligned} n!P^tW[x_1 ,..., x_n] &= P^t\hat{W}\left(\sum_{i=1}^{n} x_i\right) - \cdots - (-1)^{n-1} \sum_{j=1}^{n} P^t\hat{W}(x_j) \\ &= P^t\hat{W}\left(\sum_{i=1}^{n} P^tx_i\right) - \cdots - (-1)_n^{-1} \sum_{j=1}^{n} P^t\hat{W}(P^tx_j) \\ &= n!P^tW[P^tx_1 ,..., P^tx_i]. \end{aligned}$$

Using this result and $P_t = I - P^t$ it follows easily that

$$P^tW[x_1 ,..., P_tx_j ,...] = P^tW[P^tx_1 ,..., 0,..., P^tx_n] = 0, \qquad j = 2,..., n.$$

Since $x_j = P^t x_j + (I - P^t)x_j$, $j = 2,..., n$, and using the multilinearity of W it follows that

$$P^t W[x_1, ..., x_j, ..., x_n] = P^t W[x_1, ..., P^t x_j, ..., x_n]$$

which completes the proof.

The same algebraic operations work with $I - P^t$ and using the definition of anticausality, Lemma 2 can be dualized. These results are all gathered in the next lemma.

LEMMA 3. *Let n-linear W generate n-power $\hat{W}$. Then $\hat{W}$ is causal (anticausal, memoryless) if and only if W is causal (anticausal, memoryless) in each argument.*

A causal map f is prestrictly causal if for some finite mesh; $\Delta_i = P^i - P^{i-1}$, such that $\sum_{i=1}^{N} \Delta_i = I$,

$$f = \sum_{i=1}^{N} \Delta_i f P^{i-1}. \tag{3}$$

LEMMA 4. *Let symmetric n-linear W generate n-power $\hat{W}$. If $\hat{W}$ is prestrictly causal then $[x_1, ..., x_n]$ is prestrictly causal in all variables.*

Proof. Since $\Delta_i \Delta_j = 0 : i \neq j$ it follows that Eq. (3) holds if and only if $\Delta_i f = \Delta_i f P^{i-1}$, $i = 1,..., N$. The lemma follows then by an obvious modification of the proof of Lemma 3 above.

Our attention turns now to the strictly causal class of functions. A sequence, $\{\hat{W}_n\}$, of n-power operators converges *uniformly* to a function f if for arbitrary $\epsilon > 0$ there exists N such that $\| f(x) - \hat{W}_\alpha(x) \| \leqslant \epsilon \| x \|^n$ all $x \in H$, $\alpha > N$. The sequence $\{W_\alpha\}$ converges *strongly* to f if for every $x \in H$ and $\epsilon > 0$ there exists N such that $\| f(x) - W_\alpha(x) \| \leqslant \epsilon$ all $\alpha > N$. The strictly causal class changes with the type of convergence but fortunately the next results remain the same. To simplify matters we focus on the bipower case, the n-power proof, using Eq. (1), being a transparant modification.

LEMMA 5. *Let $\{\hat{W}_\alpha\}$ be a sequence of bipower operators which converges uniformly (strongly) to bipower $\hat{W}_0$. If W_α and W_0 are the respective bilinear symmetric generators then W_α converges to W_0 uniformly (strongly).*

Proof (Strong Closure). For arbitrary $x_1, x_2 \in H$ and $\epsilon > 0$ pick N such that $\| \hat{W}_\alpha(z) - \hat{W}_0(z) \| \leqslant 2\epsilon/3$ for $\alpha > N$ at the three points $z = x_1, x_2, x_1 + x_2$. Then using norm inequalities on the identity

$$2\{W_\alpha[x_1, x_2] - W_0[x_1, x_2]\} = \{\hat{W}_\alpha(x_1 + x_2) - \hat{W}_0(x_1 + x_2)\} \\ -\{\hat{W}_\alpha(x_1) - \hat{W}_0(x_1)\} - \{\hat{W}_\alpha(x_2) - \hat{W}_\alpha(x_2)\} \tag{4}$$

we have

$$\| W_\alpha[x_1, x_2] - W_0[x_1, x_2] \| \leqslant \epsilon, \qquad \alpha > N.$$

(*Uniform Closure*). Suppose now that for every $\epsilon > 0$, N exists such that $\| W_\alpha(x) - W_0(x)\| \leqslant \epsilon \| x \|^2$, $x \in H$. Using norm inequalities on Eq. (4) we have

$$\| W_\alpha[x_1, x_2] - W_0[x_1, x_2]\| \leqslant \epsilon/2\{\| x_1 + x_2 \|^2 + \| x_1 \|^2 + \| x_2 \|^2\}.$$

and hence

$$\sup_{\|x_1\|=\|x_2\|=1} \| W_\alpha[x_1, x_2] - W_0[x_1, x_2]\| \leqslant 3\epsilon.$$

Note now that $W_\alpha - W_0$ is bilinear and hence the proof is complete.

Our final result in the present direction is the following:

LEMMA 6. *The bipower map $\hat{W}_0$ is strictly causal if and only if its symmetric bilinear generator is strictly causal in each argument.*

Proof. If W is strictly causal then for every $\epsilon > 0$ there exists a mesh such that

$$\hat{W}_\alpha = \sum_{j=1}^{\alpha} \Delta_j \hat{W}_0 P^{j-1}$$

satisfies

$$\| \hat{W}_\alpha(x) - \hat{W}_0(x)\| \leqslant \epsilon \| x \|^2, \qquad \text{all} \quad x \in H,\ \alpha \geqslant N,$$

or for arbitrary $x \in H$

$$\| \hat{W}_\alpha(x) - \hat{W}_0(x)\| \leqslant \epsilon',$$

respectively. Using Lemma 5 we see that

$$\| W_\alpha[x_1, x_2] - W_0[x_1, x_2]\| \leqslant 3\epsilon' \| x_1 \| \cdot \| x_2 \|.$$

Using Lemma 4 we see that $W_\alpha[x_1, x_2]$ is prestrictly causal in each argument which completes the proof.

To summarize our results, let $\{W_n, n = 0,..., N\}$ be a family of n-linear symmetric operators. The function

$$f(x) = \sum_{n=0}^{N} \hat{W}_n(x)$$

is *polynomic*.

THEOREM 1. *f is causal (memoryless, strictly causal, prestrictly causal) if and only if the symmetric generator of each n-power component is causal (memoryless, strictly causal, prestrictly causal) in all arguments.*

4. The Counterexample[1]

Let f be a continuous function on real separable H and $K \subset H$ compact. It is known [4] that, for every $\epsilon > 0$, there exists a finite polynomic function, g, such that

$$\sup_{x \in K} \| f(x) - g(x) \| \leqslant \epsilon.$$

Furthermore, it is known [1] that, in $L_2(0, \tau)$, if f is causal (strictly causal, prestrictly causal) then g can be chosen with the same property. We demonstrate here, however, that a memoryless f exists for which no memoryless approximating polynomic operator exists.

Lemma 7. *In $L_2(0, \tau)$ all bounded linear memoryless operators, T, are of the form $(Tx)(t) = f(t)\, x(t)$, $t \in (0, \tau)$, where $f(t) = (T1)(t)$.*

Proof. Let $y(t) = 1$ on $[0, \tau]$; and let

$$(Ty)(t) = f(t),\ t \in [0, \tau].$$

Consider an arbitrary step function x, that is,

$$x(t) = \sum_{i=1}^{N} c_i \chi_{\Delta_i}(t), \qquad c_i \in \mathbb{R},$$

where $\{\Delta_i\}$ is an arbitrary mesh with $\chi_{\Delta_i}(t)$ the characteristic function associated with Δ_i. Since $T = \sum_{i=1}^{N} \Delta_i T \Delta_i$, we have

$$(Tx)\,(t) = \sum_{i=1}^{N} \Delta_i T \Delta_i \left(\sum_{j=1}^{N} c_j \chi_{\Delta_j}(t) \right)$$

$$= \sum_{i=1}^{N} \Delta_i T \Delta_i c_i \chi_{\Delta_i} = \left(\sum_{i=1}^{N} c_i (\Delta_i T \chi_{\Delta_i})\,(t) \right)$$

$$= \sum_{i=1}^{N} c_i \Delta_i f(t) = f(t)\, x(t), \qquad t \in [0, \tau].$$

This is true for any step function and hence for all $L_2(0, \tau)$ since the step functions are dense.

We consider now a bilinear operator W which generates the bipower $\hat{W}$. From Theorem 1 above, $\hat{W}$ is memoryless if and only if W is memoryless in each variable. This affords a relatively simple proof of

[1] The collaboration of T. Clark in constructing this section is happily acknowledged.

LEMMA 8. *There are no nontrivial bilinear memoryless operators defined on all of* $L_2(0, \tau)$.

Proof. For arbitrary y, $W[\cdot, y]$ is linear memoryless. Let

$$f_y(t) = (W[1, y])(t), \qquad t \in [0, \tau],$$

then using Lemma 7

$$W[x, y](t) = f_y(t)\, x(t), \qquad t \in [0, \tau]. \tag{5}$$

In the same vein

$$W[x, y](t) = f_x(t)\, y(t), \qquad t \in [0, \tau].$$

Now for $1\,(t) = 1$,

$$W[1, y](t) = f_y(t) \cdot 1 - f_1(t) \cdot y(t),$$

which, substituting in Eq. (5) shows that

$$W[x, y](t) = Q(t)\, x(t)\, y(t), \qquad t \in [0, \tau],$$

where $Q = W[1, 1]$. However,

$$W[x, x] = Q(t)\, x^2(t), \qquad t \in [0, \tau],$$

and since there exist points, x, on the unit ball of $L_2(0, \tau)$ whose square value is not in $L_2(0, \tau)$ the lemma follows.

Using the same syle of proof it follows easily that every memoryless n-linear operator on $L_2[0, \tau]$ is of the form

$$(W[x_1 \cdots x_n])(t) = W[1,\ldots, 1]\,(t)\, x_1(t) \cdots x_n(t), \qquad t \in [0, \tau].$$

Hence there are *no* memoryless n-linear operators defined on all of $L_2[0, \tau]$.

In view of these results it would suffice to find a compact set, $K \subset L_2(0, 1)$ and a continuous function, f, on H such that the identity plus a constant does not approximate f on K. Our counterexample is somewhat stronger than this in that no memoryless multipower operators of order $n \geqslant 2$ can exist on K let alone on H.

EXAMPLE. Consider the function $x_0(t) = t^{-1/4}$ and the sequence $\{z_n\}$, where

$$\begin{aligned} z_n(t) &= 0 \qquad && t \in [0, 1/n), \\ &= x_0^2(t), \qquad && t \in [1/n, 1]. \end{aligned}$$

We note that $z_n(t) \leqslant z_{n+1}(t)$ on [0, 1] and moreover that

$$\int_0^1 z_n(t)\, dt = \int_{1/n}^1 t^{-1/2}\, dt = 2(1 - n^{-1/2}).$$

The monotone convergence theorem then applies with the result that x_0 is square integrable and has the $L_2(0, 1)$ norm

$$\| x_0^2 \| = 2.$$

While $x_0 \in L_2(0, 1)$ note that $x_0 \notin L_4(0, 1)$, indeed

$$\lim_{\epsilon \to 0} \int_\epsilon^1 x_0^4(t)\, dt = \lim_{\epsilon \to 0} \int_\epsilon^1 dt/t = \ln(1) - \lim_{\epsilon \to 0} \ln(\epsilon) = \infty.$$

Moreover, by inspection it follows that x_0 is *not* in $L_{2m}(0, 1)$ for any $m = 3, 4, \ldots$.

Consider now the set K defined by

$$K = \{x_r : x_r(t) = | t - r |^{-1/4}, \qquad t, r, t - r \in [0, 1]\}.$$

Each x_r is a translate of x_0 and hence $K \subset L_2(0, 1)$. K is bounded, in fact $\| x_r \|^2 \leqslant 2(2)^{1/2}$ for every $x_r \in K$. From the above discussion it is also evident that $K \cap L_{2m}(0, 1) = \varphi$ for $m = 2, 3, \ldots$. We shall demonstrate first that K is compact.

A well-known fact (see [4]) for sequences $\{f_n\}$ in L_p is that: If $f_n \to f$ a.e. and $\| f_n \| \to \| f \| < \infty$ then $\| f_n - f \| \to 0$. So let $\{f_k\}$ be an arbitrary sequence in K. Then $\{k\}$ is a sequence in compact [0, 1] and hence there exists a convergent subsequence $\{k'\}$ with limit $\sigma \in [0, 1]$. This gives a subsequence $\{f_{k'}\}$ and potential limit f_σ. It is easily shown that $\| f_{k'} \|^2 \to \| f_\sigma \|$. A sketch of the functions $f_{k'}$ and f_r shows that $f_{k'}$ converges pointwise to f_σ. Thus the arbitrary sequence $\{f_r\}$ has a convergent subsequence and K is compact.

We demonstrate next that no memoryless multipower operators, $n \geqslant 2$ exist on K. Let S denote the step functions supported on [0, 1]. Since S is dense in $L_2(0, 1)$ it suffices to consider a bipower map

$$W(x)(t) = \mu(t)\, x^2(t), \qquad t \in [0, 1],$$

where $\mu \in S$.

If $\mu \neq 0$ then there exists at least one interval $\Delta_i \subset [0, 1]$ such that measure $(\Delta_i) > 0$ and $\mu(t) \neq 0$, $t \in \Delta_i$. Let $r \in \Delta_i$ and note that

$$|[W(x_r)](t)|^2 = \mu^2(r)\, x_r^4(t), \qquad t \in \Delta_i .$$

However, by our earlier computation $x_r \notin L_4(\Delta_i)$ hence $\| W(x_r)\| = \infty$ or W is not well defined on x_r and hence K.

This completes the counterexample and our study.

REFERENCES

1. W. A. PORTER, T. M. CLARK, AND R. M. DESANTIS, Causality structure and the Weierstrass theorem, *J. Math. Anal. Appl.* **52** (1975), 351–363.
2. R. M. DESANTIS AND W. A. PORTER, On the Analysis of Feedback Systems with a Polynomic Plant, *Internat. J. Control* **20** (1974).
3. W. P. VOVINGER, Mean Convergence in L^p Spaces, *Proc. Amer. Math. Soc.* **34** (1972).
4. P. M. PRENTER, On polynomial operators and equations, *in* "Nonlinear Functional Analysis and Applications" (L. B. Rall, Ed.), Academic Press, New York, 1971.
5. W. A. PORTER, Data interpolation, causality structure and system identification, *Inform. Contr.* **29** (1975), 217–233.
6. M. SCHETZEN, Measurement of the kernels of a nonlinear system of finite order, *Internal. J. Control* **1** (1965).