2004

Reconstructability Analysis with Fourier Transforms

Martin Zwick
Portland State University, zwick@pdx.edu

Citation Details
Martin Zwick, (2004) "Reconstructability analysis with Fourier transforms", Kybernetes, Vol. 33, No.: 5/6, pp. 1026 - 1040

RECONSTRUCTABILITY ANALYSIS WITH FOURIER TRANSFORMS

Martin Zwick
Systems Science Ph.D. Program, Portland State University, Portland, OR 97207-0751
zwickm@pdx.edu http://www.sysc.pdx.edu/Faculty/Zwick

KEYWORDS: reconstructability analysis, Fourier transforms, information theory, state-based modeling, k-systems analysis, log-linear modeling

ABSTRACT

Fourier methods used in 2- and 3-dimensional image reconstruction can be used also in reconstructability analysis (RA). These methods maximize a variance-type measure instead of information-theoretic uncertainty, but the two measures are roughly colinear and the Fourier approach yields results close to those of standard RA. The Fourier method, however, does not require iterative calculations for models with loops. Moreover the error in Fourier RA models can be assessed without actually generating the full probability distributions of the models; calculations scale with the size of the data rather than the state space. State-based modeling using the Fourier approach is also readily implemented. Fourier methods may thus enhance the power of RA for data analysis and data mining.

I. INTRODUCTION

In reconstructability analysis (RA) applied to probabilistic systems, probability distributions for subsets of variables specified by a model are joined together to calculate a probability distribution for the full set of variables (Klir 1985; Krippendorff 1986). Similarly, in image reconstruction (IR) used in electron microscopy, tomography, and other areas, lower-dimensionality projections are combined to yield a full-dimensionality density function (Zwick & Zeitler, 1973). This paper will show that Fourier techniques used in IR can be applied also to RA (and thus to log-linear modeling (Bishop et al 1978; Knoke & Burke 1980) which closely resembles RA).

There are important differences between IR and RA. IR treats continous density functions defined on interval scale variables in 2 or 3 spatial dimensions. Projections arise from rotations of the object or of the imaging source, and are not mutually orthogonal. RA, by contrast, considers probability distributions, defined on a discrete, in fact nominal, domain of higher dimensionality, and the projections are all mutually orthogonal. Despite these differences, essentially the same task, namely *composition* of lower-dimensional projections to obtain a higher dimensional function, is accomplished in both areas.

The Fourier method used in IR is as follows. Since the Fourier transform of a projection of a distribution is a central section (a section passing through the origin) of the transform of the function, measured projections can be combined by calculating their transforms, collecting these sections together in Fourier space, and doing an inverse transform to obtain a function which has these projections.

For a compact review of reconstructability analysis, see (Zwick, 2001). RA comprises two problem types: reconstruction and identification. Identification is the simpler of the two and closely resembles IR: the task is composition of a set of projections into a higher dimensionality distribution. This is done by the "iterative proportional fitting" (IPF) algorithm, in which projections are sequentially imposed on a calculated distribution initialized as uniform. Iterations of such impositions eventually converge on a distribution consistent with all projections. Actually IPF is needed only for models with loops, since for models without loops algebraic (non-iterative) solutions are available. But most models, and virtually all complex models, have loops.

Reconstruction, however, is the problem most commonly encountered in RA and is the focus of this paper. The task here is to represent and approximate a distribution with a set of its lower dimensional projections. Reconstruction thus consists of three steps: (a) projection, (b) composition, and (c) evaluation. Projection yields the lower dimensionality distributions whose adequacy is being explored. Composition is done as in identification. Evaluation assesses the difference between the computed IPF distribution and the observed distribution.

The projection and evaluation steps of RA do not pose serious computational problems, as they scale with the size of the data and not the state space. It is the *composition* step which poses the primary challenge, and this challenge is two-fold: (i) Many iterations are sometimes needed for IPF to converge, and (ii) IPF calculates probabilities for the entire state space, even when data are sparse. The computer time and space requirements of IPF restricts the applicability of RA by severely limiting the number and cardinalities of the variables which can be considered.

What motivates this paper is the observation that in IR, composition is accomplished in a *single* iteration: one

simply takes the Fourier transforms of all projections, collects together the resulting sections in frequency space, and performs the inverse Fourier transform. (Computations are done efficiently by using the fast Fourier transform.) If such a single-iteration method for composing projections were available in RA, it would enhance the power of RA for exploratory modeling. It will be shown below that this is indeed possible. Specifically, IR-type composition provides a single-iteration approximation to IPF, which can be used for rough searches through the lattice of possible models.

This addresses only the first of the two difficulties posed by IPF, since Fourier composition also involves the entire state space. It turns out, however, that *back projection*, a procedure equivalent to the Fourier approach, allows a "reduced" composition step to be done that calculates probabilities only for *observed* states. If the IR approximation to standard RA is adequate, exploratory modeling with the IR approach can thus bypass both the time and space limitations of IPF.

The IR approach also allows the easy implementation of state-based modeling, a variant of RA pioneered by Jones (1985a,b) and currently under further development (Johnson and Zwick 2000; Zwick and Johnson 2002). This implementation however still scales with the state space and not the data.

II. A SIMPLE 2-D FOURIER RA EXAMPLE

To investigate whether the IR Fourier approach might be applied to RA, consider the 2-dimensional RA problem, for which there are only two possible models: AB, the "saturated" model (the data), and A:B, the "independence model." The simplest case occurs where variables are dichotomous (binary) and the AB distribution is a 2x2 contingency table, as illustrated in Table 1(a). This table requires 3 parameter values for its specification, which is the *degrees of freedom* (df) of AB; this is suggested by the shading of 3 (arbitrarily chosen) cells.

Table 1. (a) Observed probability distribution, AB, and (b) independence model, A:B

		B 0	B 1	
A	0	.1	.2	.3
	1	.3	.4	.7
		.4	.6	

(a)

		B 0	B 1	
	0	.12	.18	.3
	1	.28	.42	.7
		.4	.6	

(b)

The independence model, A:B, is the distribution, which is the product of the margins of AB, as shown in Table 1(b). Only 2 parameters (arbitrarily chosen and shown shaded), one in each projection, are needed to specify this model.

The A:B distribution is the solution to the maximization of information-theoretic uncertainty subject to model constraints, i.e., to the problem:

maximize $U = - \Sigma\Sigma\, q(i,j) \log q(i,j)$ subject to

$$\Sigma\, q(0,j) = p(0,\bullet) = .3 \qquad \Sigma\, q(1,j) = p(1,\bullet) = .7$$
$$\Sigma\, q(i,0) = p(\bullet,0) = .4 \qquad \Sigma\, q(i,1) = p(\bullet,1) = .6$$
$$\Sigma\Sigma\, q(i,j) = 1$$

where $p(j,\bullet) = \Sigma_k\, p(j,k)$ and $p(\bullet,k) = \Sigma_j\, p(j,k)$, and where p and q refer to the observed (AB) and calculated (A:B) distributions, respectively. Although there are four projection equations, given the fifth equation which sets the sum of the probabilities to 1, there are only two linearly independent equations of constraint (one for each projection) hence df(A:B)=2. As noted above, Table 1(b) is the maximum uncertainty distribution subject to the constraints of the A:B model. It can also be generated directly as an algebraic function (here, the simple product) of the projections of the original distribution of Table 1(a). In general, however, and specifically for models with loops, one cannot derive the model distribution algebraically, and the IPF algorithm must be used.

The IR Fourier method is applied to this problem as follows. The discrete 2-dimensional Fourier transform of p(j, k), where $j=1,2,...N_j$ and $k=1,2,...N_k$) and where(J,K) are the indices in Fourier space corresponding to (j,k), is:

$$P(J, K) = \Sigma_j\, \Sigma_k\, p(j, k) \exp [\ 2\pi i\ (\ j\ J / N_j\ + k\ K / N_k\ )\ ]$$

Extension to higher dimensions is straightforward. From the theorem that the Fourier transform of a projection is a central section, the projections, p(j,•) and p(•,k) have for their Fourier transforms the central sections P(J,0) and P(0,K), respectively, as follows.

Equation(1) *Calculating central sections from projections*

$$P(J,0) = \Sigma\, p(j,\bullet)\ \exp [\ 2\pi i\ (\ j\ J\ / N_j\ )\ ]$$
$$P(0,K) = \Sigma\, p(\bullet,k) \exp [\ 2\pi i\ (\ k\ K / N_k\ )\ ]$$

The Fourier transform of $p(j,\bullet) = \{.3, .7\}$ is $\{1.0, -0.4\}$, since

$$P(0,0) = .3 \exp [2\pi i\ (0)(0)/\ 2]\ + .7\ [\exp\ 2\pi i\ (1)(0)/2] = 1.0$$
$$P(1,0) = .3 \exp [2\pi i\ (0)(1)/\ 2]\ + .7\ [\exp\ 2\pi i\ (1)(1)/2] = -0.4$$

and the transform of $p(\bullet,k) = \{.4, .6\}$ is $P(0,K) = \{1.0, -0.2\}$. If one collects together these two central sections, one has in Fourier space the transform shown in Table 2 (the central sections are shaded).

Table 2. Central Fourier sections of Table 1(a)

	0	*1*
0	*1.0*	*-0.2*
1	*-0.4*	*0.0*

In the Fourier approach to RA, after collecting together the central sections dictated by the model, one then does an *inverse* Fourier transform to obtain the q distribution corresponding to these projections, as follows:

Equation (2) *Inverse transform of set of central sections*:

$$q(j,k)= \Sigma\Sigma[P(J,0)+P(0,K)]\exp[-2\pi i\,(jJ/N_j+kK/N_k)] \quad - P(0,0)$$

It would be natural to presume that the inverse transform of Table 2 would yield the independence model distribution of Table 1(b). The origin term, 1.0, of the Fourier distribution corresponds to $\Sigma\Sigma\ p(j,k) = 1$. By itself, this term generates the uniform distribution. If one adds to this term the sections corresponding to the two projections, one might expect the result to be the maximally uniform distribution subject to the projections as constraints, i.e., Table 1(b).

This expectation is *not* correct. The inverse transform of Table 2 is Table 1(a) and not Table 1(b). Table 2 is actually the *full* transform of Table 1(a), i.e., the *non*-central part of the transform of Table 1(a) is 0. Table 1(a), the AB distribution which in standard RA exhibits non-zero constraint is fully specified in the Fourier approach by its projections. By contrast, the Fourier transform of Table 1(b), the independence model in standard RA, is shown in Table 3. Its non-central Fourier coefficient is non-zero (0.08). Thus, Table 2, the transform of a distribution *with* constraint seems to have df=2, while Table 3, the transform of a distribution *without* constraint seems to have df=3.

Table 3. Fourier transform of Table 1(b)

	0	*1*
0	*1.0*	*-0.2*
1	*-0.4*	*0.08*

This anomaly is the result of the particular distribution chosen for analysis. This example was in fact chosen to highlight the differences between the Fourier approach and standard RA. Because the Fourier transform is a linear operation, the inverse transform of the collected sections yields, not the *product* of the projections as in standard RA,

$$q(j,k)_{\text{standard RA}} = p(j,\bullet)\ p(\bullet,k)$$

but a scaled *sum* of them, sometimes referred to as the operation of "back-projection" (BP)

Equation (3) *Back-projection*:

$$q(j,k)_{\text{Fourier RA}} = p(j,\bullet) / N_j \ + \ p(\bullet,k) / N_k \ - \ 1 / N_jN_k$$

A projection B projection origin correction

Back projection is equivalent to implementing Equations (1) and (2) in one step by operating on the projections directly without actually doing any transform. Note that if data were sparse, this equation could be used to evaluate q(j,k) only for those (j,k) which were actually observed; the entire state space of (j,k) would not have to be generated. Applying BP to Table 2 yields Table 1(a), as follows:

Table 4. Fourier composition of Table 1(a) from projections using BP (Equation (3))

.1	.2	=	.15	.15	+	.20	.30	-	.25	.25
.3	.4		.35	.35		.20	.30		.25	.25
			$p(j,\bullet) / N_j$			$p(\bullet,k) / N_k$			$1 / N_jN_k$	

By contrast, the inverse transform of Table 3 yields Table 1(b) because of the extra contribution of the non-central 4th term (0.08) to the calculated distribution. This contribution,

		B	
		0	1
A	0	+.02	-.02
	1	-.02	+.02

when added to Table 1(a), gives Table 1(b). Consider Table 5, which has the same projections as Table 1(a) and (b). Fourier reconstruction of Table 5 also yields Table 1(a).

Table 5. A 2nd distribution with identical margins

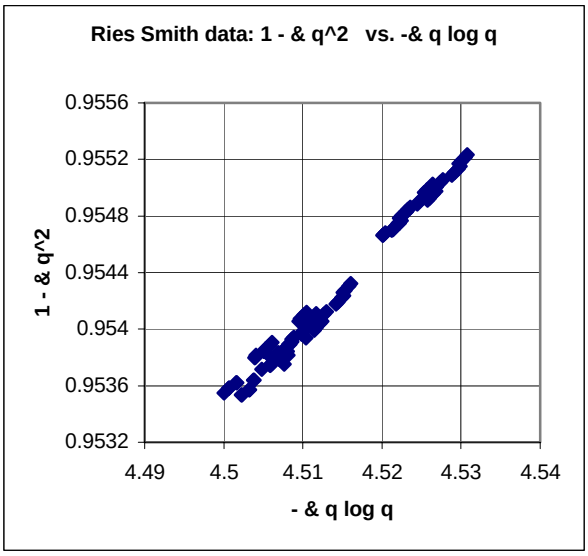
		B		
		0	1	
A	0	.30	.00	.3
	1	.10	.60	.7
		.4	.6	

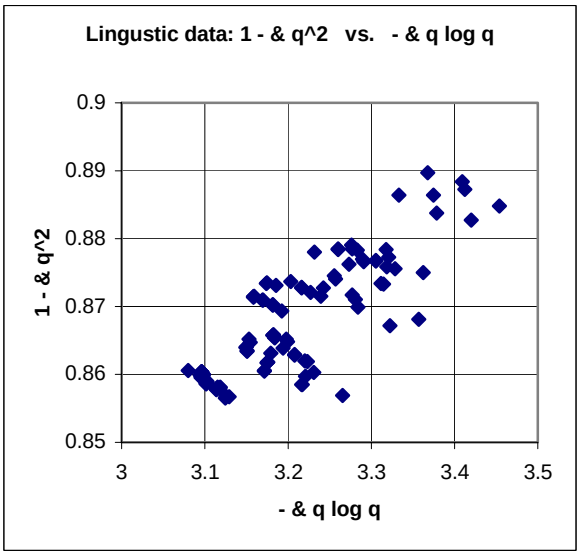
The point of all this is that Table 1(a) and not Table 1(b) is the A:B "independence model" for the *Fourier approach*. The reason the Fourier method gives results different from standard RA is that while RA composition maximizes the uncertainty, $-\Sigma\ q \log q$, the Fourier approach minimizes Σq^2 , or, equivalently maximizes, $-\Sigma\ q^2$. Table 1(a) is the independence model in the Fourier approach because it is the maximum $-\Sigma\ q^2$ distribution for the {.3, .7} and {.4, .6} margins. Because of this, it is the reconstructed distribution when the data is Table 1(a) or Table 1(b) or Table 5.

That the Fourier method minimizes $\Sigma\ q^2$ subject to the model constraints can be seen more directly from the fact that, for Q the Fourier transform of q (and with proper

scaling), $\Sigma |Q|^2 = \Sigma q^2$. The model constraints give Q the central sections of P, which are the transforms of the projections included in the model, with all other coefficients in Q being 0. Q thus embodies nothing beyond the model constraints and thus generates the minimum Σq^2.

Figure 1. 1 - Σq^2 is roughly linear with - Σ q log q. The measures are plotted for all 114 models for the Ries Smith data (top) and linguistic data (bottom).





Although the Fourier approach does not maximize -Σq log q it might be usable for RA because the - Σq^2 it maximizes is roughly colinear with uncertainty. This shown in Figure 1 above which plots 1 - Σq^2 vs. - Σ q log q for all possible 4-variable models for the Ries-Smith marketing data (Bishop et al 1978) and some linguistic data currently under investigation (Zwick and McCall 2002).

A linear relationship is evident for both data sets, although the relationship is obviously stronger for the first data set (r=.99) than for the second (r=.84). The generality and the factors which affect the strength of this colinearity need to be investigated further, both empirically and analytically. However, there are reasons to expect that this result is robust since 1 - Σq^2 is used in economics and ecology as an alternative to uncertainty to quantify diversity. In economics the sum-squared measure is known as the Herfindahl index (Jacquemin & Berry, 1979). Intuitively, if the total probability of 1 is divided into many small terms, the sum of their squares will be small, so 1 - Σq^2 is a plausible measure of diversity (uncertainty). Because of the rough colinearity of - Σ q log q and - Σq^2, maximizing the latter expression is likely to give results close to maximizing the former expression when both maximizations have the same constraints. So the Fourier approach has distinct promise for RA.

III. METHODOLOGY OF FOURIER RA

Structure Specification

So far only a 2-dimensional example has been discussed. Consider the case of three variables, where the existence of structures (sets of relations) first arises, and where the Lattice of Relations and the Lattice of Structure are shown in Figure 2 and Figure 3, respectively. (A "model" might be defined as a "structure" applied to data, but the distinction is subtle and this paper does not always insist on it.)

Figure 2. Lattice of Relations for a 3-variable system

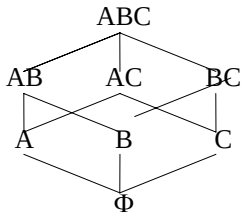
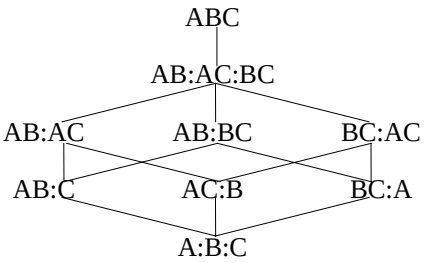


Figure 3. Lattice of Structures for 3-variable system



The Fourier model of ABC is its full 3-dimensional transform. Next in the Lattice of Structures is AB:AC:BC. Its transform, Q, consists of sections, which are the

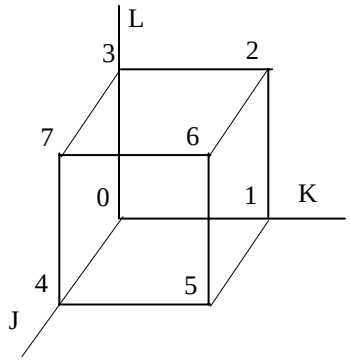
transforms of the AB, AC, and BC projections. If the AB projection is written as p(A,B) and its Fourier transform as P(A,B), then the central section $Q_{AB:AC:BC}(A,B,0) = P(A,B)$ contains the information on AB. Similarly, the sections $Q_{AB:AC:BC}(A,0,C) = P(A,C)$ and $Q_{AB:AC:BC}(0,B,C) = P(B,C)$ contain the information on the AC and BC projections. Thus $Q_{AB:AC:BC}(A,B,C)$ consists of a *subset* of coefficients from the full P(A,B,C) transform, namely those for which either A=0 *or* B=0 *or* C=0. Thus

$$Q_{AB:AC:BC}(A,B,C) = \{P(A,B,0)\} \cup \{P(A,0,C)\} \cup \{P(0,B,C)\}.$$

One can list the Fourier components for a model in terms of a "dual," which indicates which central sections are to be included. For example, the dual of AB:AC:BC is *C:B:A*, written in italics, which summarizes the condition that this model includes Fourier coefficients where C=0 or B=0 or A=0. (The colon in *C:B:A* means this inclusive "or".) Applying this to the full Lattice of Structures yields Table 6. The table is read as follows: a coefficient P(A,B,C) is included in a structure if the indicated zero condition defined by the dual holds for that coefficient.

To show this in greater detail, Figure 4 represents P(J,K,L), the Fourier transform of the ABC distribution, p(j,k,l). Using the labels from this figure, the Fourier coefficients included in all structures are tabulated in Table 6. The degrees of freedom for these structures are also listed.

Figure 4. Fourier coefficients as model parameters. The numbers 0 to 7 label the Fourier coefficients, P(J,K,L), where variables are binary, i.e., J=0,1 and similarly for K and L. For example, point 6 is P(1,1,1) and point 0 is P(0,0,0). P(0,K,L), the central section which transforms p(•,k,l), contains points 0,1,2,3.



Fourier reconstruction is additive in a way that conventional RA is not. A model is a set of relations. If relation 1 has coefficient set s_1 and relation 2 has coefficient set s_2, then a model including both relations has coefficient set $s_1 \cup s_2$. By virtue of this additivity models with loops do not require an iterative procedure to derive the calculated distribution.

Table 6. Conditions for inclusion of P(A,B,C) coefficients in models. A variable in the *dual* must be 0 for a coefficient to be included in the model.

				coefficients included							
level	structure	*dual*	df	0	1	2	3	4	5	6	7
0	ABC	Φ	7	+	+	+	+	+	+	+	+
1	AB:AC:BC	*C:B:A*	6	+	+	+	+	+	+		+
2.1	AB:AC	*C:B*	5	+	+		+	+	+		+
2.2	AB:BC	*C:A*	5	+	+	+	+	+	+		
2.3	AC:BC	*B:A*	5	+	+	+	+	+			+
3.1	AB:C	*C:AB*	4	+	+			+	+		
3.2	AC:B	*B:AC*	4	+			+	+			+
3.3	BC:A	*A:BC*	4	+	+	+	+				
4.1	A:B:C	*BC:AC:AB*	3	+	+		+	+			

Degrees of Freedom

Table 6 suggests that the degrees of freedom of a Fourier model is the number of coefficients in all of the central sections defined by the model minus 1, to omit the origin term which corresponds to the sample size for frequency distributions or to the sum of probabilities being 1. Actually, the calculation of df is a little more complicated. Fourier coefficients are in general complex, which suggests that the real and imaginary parts should contribute 2 degrees of freedom for each coefficient. However, coefficients for transforms of real functions come in conjugate-symmetric pairs, where P(J) = P*(-J), where * means complex conjugate and J is in general a vector, so a *pair* of coefficients contributes 2 degrees of freedom. When J = -J coefficients occur in singletons, which because of conjugate symmetry must be real, again contributing 1 degree of freedom per coefficient. In Table 6 above, all coefficients are singletons, so the df calculation is trivial, but in general both pairs and singletons will occur.

Model Error

When the Fourier approach is used as an alternative RA framework, modeling any specific structure generates zero conditions from the dual of the structure. These allow one to construct Q, the transform of the model, which, inverse-transformed, yields q. q can be assessed using either (a) the standard RA transmission, $T(q) = \Sigma\, p \log p/q$ or (b) the sum-squared-error $SSE = \Sigma\,(p - q)^2 = \Sigma\,(P - Q)^2$, an error measure more naturally associated with the Fourier method. Only the coefficients absent in the model generate error, i.e., $SSE = \Sigma\,(P - Q)^2 = \Sigma_{\text{omitted in Q}}\, P^2$. This means that SSE can be evaluated *without generating q* by taking the sum of squares of the omitted Fourier coefficients. A Fourier-based RA search, using SSE to evaluate models, does not need to inverse transform Q into q.

Further, since contributions of missing coefficients to the error are mutually independent one from another, these errors do not have to be generated anew for every different model. Instead, errors can be calculated and stored for every *relation* in the Lattice of Relations. The error in any model can then be generated algebraically from these stored relation-SSEs. Refer again to Figure 4 where the numbers 0-7 represents the Fourier coefficients, P. The model AB:AC:BC includes all coefficients where A=0 or B=0 or C=0; its error is thus represented above by "6", i.e., P(1,1,1). This can be derived from the errors of the model's component relations as follows:

SSE(AB:AC:BC) = SSE(AB) +SSE(AC) +SSE(BC)
-SSE(A)-SSE(B)-SSE(C) +SSE(Φ)

=+(2 3 6 7)+(1 2 5 6)+(4 5 6 7)
-(1 2 3 5 6 7)-(2 3 4 5 6 7)-(1 2 4 5 6 7)+(1 2 3 4 5 6 7) = 6

where Φ is the uniform distribution model, generated from only the origin coefficient of the transform. It seems likely (no proof is offered here), for relations in a model written as R_1, R_2, ..., where $R_j \cap R_k$ is the relation defined by variables common to R_j and R_k, that SSE can be written as follows:

$$SSE(R_1:R_2:...) = \Sigma\, SSE(R_j) - \Sigma\Sigma\, SSE(R_j \cap R_k) + \Sigma\Sigma\Sigma\, SSE(R_j \cap R_k \cap R_l) - ...$$

Model information computed from SSE is closely related to transmission-defined information, as follows:

$I_{Fourier\ RA}$ = [SSE(A:B:C:D) - SSE(model)] / SSE(A:B:C:D)
$I_{standard\ RA}$ = [T(A:B:C:D) - T(model)] / T(A:B:C:D)

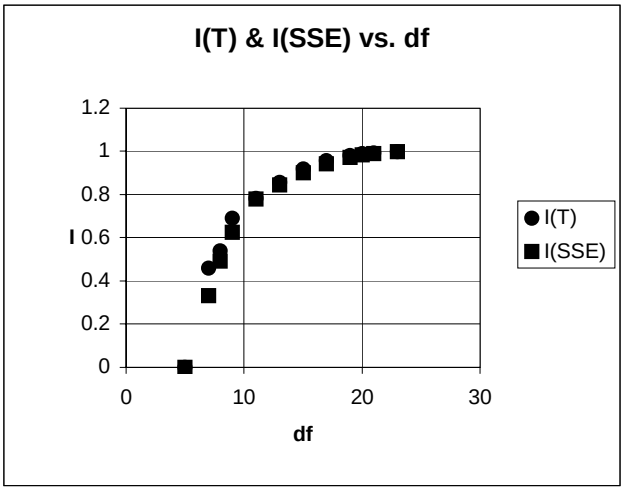
This is shown in Figure 5 which is based on the data used for Figure 1(top). The figure shows, for every value of df, $I_{standard\ RA}$ (circles) and $I_{Fourier\ RA}$ (squares) for the highest information model using standard RA. The Fourier results approximate the standard results, especially at high information. This plot is closely related to Figure 1, since T is linear with U and SSE with Σq^2:

T(model) = U(model) – U(data)
SSE(model) = $\Sigma_{omitted\ in\ Q}$ $P^2 = \Sigma p^2 - \Sigma q^2$(model)

Sparse data

Defining parameters in Fourier space is an approach to RA model construction whose practicality depends on the data. In all the examples considered in this paper, the contingency table is full, in that each cell has a frequency greater than 1. This accords with the Chi-square rule of thumb that the sample size should ideally be at least about 5 times the number of states. Where the data is *sparse*, however, Fourier transformation will spread the data throughout Fourier space. For example, the transform of a gaussian is a gaussian, and if the gaussian gets narrower in distribution space, it gets broader in Fourier space. Thus, Fourier representation of sparse data is likely to have higher df, which may defeat the goal of compression. This issue is being investigated further. Possibly, wavelet, as opposed to Fourier, transforms might be an alternative approach to modeling sparse data, since for sparse data, the global character of Fourier transforms may be disadvantageous, while the local character of wavelet transforms may be useful. A wavelet approach to sparse data may need variable states to be relabeled and thus reordered to concentrate the data locally.

Figure 5. Information of models using standard and Fourier RA (Ries Smith data) (A:B:C:D has df=5.)



Back Projection as an Alternative to IPF

All this presumes that the Fourier coefficients, P, are the RA model parameters and that composition is done with Equation (2). However, a "reduced" composition can be done with Back-Projection Equation (3), which operates on distributions (not Fourier coefficients). If one wants only to screen models by evaluating T and thus needs only q values for observed states, this can be done with BP, which approximates IPF in a single iteration and, used for this purpose, scales with the data, not the state space.

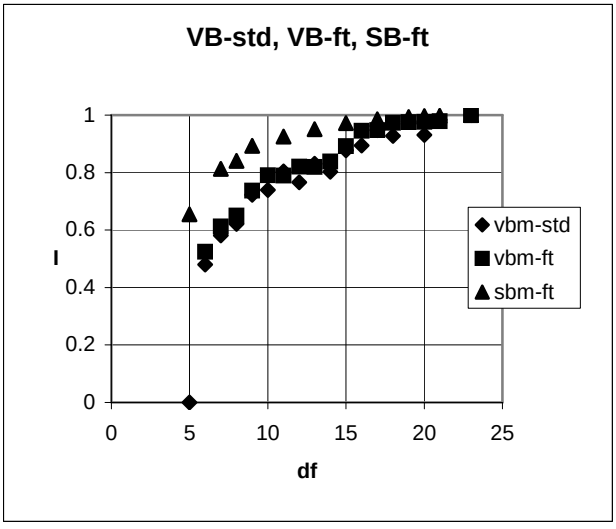
Note that Equation (3) can yield negative q values. If the Fourier approach is used in RA only for model evaluations in exploratory searches, and not as a source of full q distributions, this may not be a problem; also, correctives are imaginable. Still, this possibility is one which requires further theoretical and computational exploration.

IV. STATE-BASED MODELING

The Fourier components need not be restricted to central sections. One could choose, for a df = n model, the n *biggest* Fourier coefficients from the original transform. This amounts to the Fourier equivalent of the "state-based" modeling approach (Johnson and Zwick 2000; Zwick and Johnson 2002) derived from the "k-systems analysis" of Jones (1985a,b). In state-based, as opposed to variable-based, modeling, an RA model does not need to be defined in terms of *complete* projections (margins), but can instead be defined in terms of the probabilities of an arbitrary set of states (as long as the probabilities are linearly independent). Applying this notion to the Fourier approach to RA, models need not consist only of central sections but can be any set of Fourier coefficients.

State-based modeling has a Lattice of Structures enormously greater than variable-based RA. This poses the problem of how to search this lattice. Jones (1985c) proposed a path-dependent procedure: one selects the most information-rich state, the 2nd most information-rich *given* the prior selection of the 1st state, and so on. Because of the path dependence of this algorithm, one cannot be sure that a state-based model, involving n states actually consists of the n most informative states. This uncertainty disappears in the Fourier approach to state-based modeling in which a model is defined, not from central sections but by selecting Fourier coefficients from anywhere in the transform in descending order of magnitude. For a model with df = n, one simply selects the n biggest coefficients. The information content of such a state-based model will always be equal or superior to a df = n variable-based RA model.

Figure 6. Fourier transform variable- and state-based models, compared to standard variable-based models. (Ries Smith data)



This is shown in Figure 6 which shows, for every df, the I for standard variable-based RA for the highest information model (diamonds). On the same scale, I from transmissions of variable- and state-based models using Fourier transforms are also plotted. Clearly, $I_{VB\text{-}ft} \approx I_{VB\text{-}std}$ but $I_{SB\text{-}ft} > I_{VB\text{-}std}$.

While state-based modeling achieves greater compression than variable-based modeling, it has the disadvantage that Fourier coefficients are not as interpretable as a distribution and its projections. This is especially so since the values of these coefficients depends upon the specific ordering of the states of the variables, but this ordering is arbitrary.

The Fourier approach thus constitutes an alternative framework for doing state-based RA. State-based RA, like variable-based RA, can also use Fourier ideas just by using Equation (3) of BP as an efficient approximation to IPF.

V. SUMMARY

This paper demonstrates that the Fourier approach of Image Reconstruction can be applied to Reconstructability Analysis. Although Fourier reconstruction maximizes - Σq^2, while standard RA maximizes - Σ q log q, the two objective functions are roughly colinear. The Fourier approach can be used in a variety of ways:

- This approach provides an alternative framework for RA. Projections can be collected in Fourier space and composition done in a single inverse transform. Calculated distributions can be evaluated with the standard T measure.

- The search through the Lattice of Structures does not need to generate model distributions. If SSE is used instead of T, model error can be assessed without inverse transformation by computing the error resulting from the omitted transform coefficients. This can be generated directly from the set of relations included in the model.

- One can use the Fourier approach more narrowly by simply replacing IPF with BP which is non-iterative and, when used for the evaluation of a model statistic like T, does not rquire operations on the whole state space. (Using BP to obtain a full q distribution, however, scales with the state space.)

- State-based models, which can capture more information than variable based models of the same df, are also easily implemented. There is no path-dependence in Fourier state-based modeling, and the best model can easily be selected for every df value. Alternatively, state based modeling can be

done in the usual way, but models might be fitted by BP, as an efficient approximation to IPF.

Only a proof-of-concept of the Fourier approach to RA is provided here. Theoretical and computational issues are still being explored. This project is part of a larger effort in "discrete multivariate modeling" (Zwick, 2002), i.e., RA, which includes software development Willett and Zwick 2002) that will eventually encompass the Fourier approach. Work so far, however, has demonstrated clearly that reconstructability analysis can be approached with Fourier techniques. This is not surprising. Walsh, Haar, and other transforms are routinely used in logic design and machine learning, and methods in these fields overlap set-theoretic (crisp possibilistic) reconstructability analysis. Roughly speaking, the use proposed here of Fourier transforms in probabilistic RA parallels the use of the above discrete transforms in crisp possibilistic RA.

VI. REFERENCES

Bishop, Y. M.; S.E. Feinberg; and P.W. Holland. 1978. *Discrete Multivariate Analysis*. MIT Press, Cambridge.

International Journal of General Systems (IJGS) Special Issue on GSPS 1996, vol. **24** pp. 1-2.

Jacquemin, A. P. and C.H. Berry. 1979. "Entropy Measure of Diversification and Corporate Growth." *The Journal of Industrial Economics*, vol. 27 (4), pp. 359-369.

Johnson, M.S. and M. Zwick. 2000. "State-Based Reconstructability Modeling for Decision Analysis." In: *Proceedings of World Congress of the Systems Sciences and International Society for the Systems Sciences 2000*, Allen, J.K. and Wilby, J.M., ed., Toronto.

Jones, B. 1985a. "Reconstructability Analysis for General Functions." *Int. J. General Systems*, vol. 11, pp. 133-142.

Jones, B. 1985b. "Reconstructability Considerations With Arbitrary Data." *Int. J. General Systems*, vol. 11, pp. 143-151.

Jones, B. 1985c. "A Greedy Algorithm for a Generalization of the Reconstruction Problem." *International Journal of General Systems* **11**: 63-68.

Klir, G. 1985. *The Architecture of Systems Problem Solving*. Plenum Press, New York.

Knoke, D. and P.J. Burke. 1980. *Log-Linear Models*. (Quantitative Applications in the Social Sciences Monograph # 20). Sage, Beverly Hills.

Krippendorff, K. 1986. *Information Theory: Structural Models for Qualitative Data (Quantitative Applications in the Social Sciences #62)*. Sage, Beverly Hills.

Willett, K. and M. Zwick. 2002. "A Software Architecture for Reconstructability Analysis." In: *Proceedings of 12th International World Organization of Systems and Cybernetics and 4th International Institute for General Systems Studies Workshop*, Pittsburgh.

Zwick, M. 2001. "Wholes and Parts in General Systems Methodology.". In: *The Character Concept in Evolutionary Biology*, edited by Günter Wagner. Academic Press, New York, pp. 237-256.

Zwick, M. 2002. "Discrete Multivariate Modeling": http://www.sysc.pdx.edu/res_struct.html

Zwick, M. and McCall, M. 2002. "Reconstructability Analysis of Language Structure." Manuscript in preparation.

Zwick, M. and M.S. Johnson. 2002. "State-Based Reconstructability Modeling." In: *Proceedings of 12th International World Organization of Systems and Cybernetics and 4th International Institute for General Systems Studies Workshop*, Pittsburgh.

Zwick, M. and E. Zeitler. 1973. "Image Reconstruction from Projections." *Optik*, vol. 38, pp. 550-565.

ACKNOWLEDGEMENTS

I thank Michael Johnson and Ken Willett for helpful comments on this paper.

BIOGRAPHY

Martin Zwick

Martin Zwick is a Professor of Systems Science at Portland State University. Prior to taking his current position at PSU, he was a faculty member in the Department of Biophysics and Theoretical Biology at the University of Chicago, where he worked in macromolecular structure and mathematical crystallography. In the 1970's his interests shifted to systems theory and methodology. Since 1976 he has been on the faculty of the PSU Systems Science Ph.D. Program and during the years 1984-1989 he was the program Coordinator and then Director. His current research interests are in discrete multivariate modeling (reconstructability analysis), "artificial life" and theoretical/computational biology, and systems philosophy.