JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 5, 46-69 (1962)

Multidimensional Curve-Fitting with Self-Organizing Automata

GABRIEL KRON

General Electric Company, Schenectady, N.Y.

Submitted by Richard Bellman

In large-scale linear problems a one-dimensional electrical network often serves as a model, at discrete points, for a function of k independent variables. Such a discrete animated model has been generalized by the author to represent a class of large-scale nonlinear problems by introducing a sequence of linear-, planar-, cubic-, etc., up to k-dimensional networks, all interconnected into a single polyhedral structure. However, a hierarchy of multidimensional networks can no longer be energized by mere currents and voltages. A sequence of multidimensional electromagnetic waves must be propagated across the polyhedron (and its dual polyhedron), in order that the waves may satisfy Stokes' theorem, as they step across the boundaries between two different-dimensional networks. Such an animated polyhedral model can represent not only a function of k independent variables, but also all its divided differences of higher order (estimated directional derivatives along the lines, planes, cubes, etc.) all simultaneously. Furthermore, if the polyhedron and its dual are immersed into a k-dimensional region filled with stationary or moving magnetohydrodynamic plasma, the amorphous field crystallizes into a sequence of $2k$ sets of transmission networks, coupled by and energized with a large number of k-dimensional magnetohydrodynamic generators ("generalized" rotating electrical machines). Even in the absence of motion (velocity terms), the crystallized field-structure may assume a self-adaptive "oscillatory" state, in which it can represent not one, but any number of arbitrarily picked functions of k independent variables, as well as their higher order divided differences—all simultaneously and automatically—without needing any adjustment or interference by the analyst.

The resultant oscillatory polyhedron (or self-organizing "automaton") is applied in the present paper to model (curve-fit) simultaneously any number of functions of a set of nonuniformly-spaced variables. In particular, simple numerical examples are shown of estimating—by regression theory and a least-square criterion—several arbitrary functions of two independent variables at four nonuniformly-spaced points on a plane. The same oscillatory model (automaton) is used, without any change, for the highly satisfactory estimation of not one but six different arbitrarily picked functions, plus their divided differences. Numerical examples with two nonoscillatory polyhedra show that the latter can satisfactorily fit only one, or at most, only a small class of functions plus their divided differences.

Introduction

Mathematical Background of Physical Concepts

In order not to frighten the mathematician or statistician readers of this paper by the electrical engineering terminology, it should be pointed out at the beginning that it is the belief of the author that many—if not all—of the physical concepts stated in these pages can be restated in purely mathematical terminology. He believes that these concepts can be fitted somewhere into the framework of modern tensor calculus that deals with such topics as the integration theory of "exterior" differential forms [1]. During the last three decades the author has been applying conventional tensor-analysis, which is now used quite extensively in the local *field* problems of differential geometry, to global electrical *network* problems. Lately the author discovered that during the same decades Cartan, DeRham, Hodge, Whitney, Steenrod, and a host of other mathematicians have extended the theory of local "tensors in the small," to global "tensors in the large," by utilizing the multiply-connected, curved polyhedral networks of algebraic topology. These mathematicians thus apply tensorial methods to the analysis and approximate solution of *combined field and network problems*. That is also the purpose of the author (and of the present paper), except that he specializes in getting numerical answers to practical *physical* problems.

Since the author is an electrical engineer and not a mathematician, his models are electrical and not mathematical networks. He believes that his electrical network researches are somewhat equivalent to replacing some of the geometrical concepts of mathematicians with physical concepts, and to retaining others as geometry. Actually, he attempts to combine mathematics, geometry, and physics into one *engineering* tool. At the same time he is definitely anxious that his physical structures should dovetail as closely as possible into the elaborate geometrical and mathematical structures built by theoretical mathematicians.

The Discretization of Fields by "Neighborhoods"

Statisticians consider the rows of an experimental $n \times k$ data-matrix X (independent variables) as n scattered points in a k-dimensional space. The single column y of the dependent variables is viewed either as weights at the points, or as a k-dimensional surface immersed in a $k + 1$ dimensional volume. This ***netpoint*** representation of a surface is quite analogous to the manner in which mathematicians are solving multidimensional partial differential equations by the use of finite-difference methods. Electrical engineers have also been employing the netpoint representation of field problems by replacing the *difference* equations with electrical one-dimensional network analogous.

During the heyday of the ac network analyzer, the author specialized in the electrical network modeling of the partial *differential* equations (not difference equations) of mathematical physics (both classical and quantum physics). He realized even then that the use of electrical networks implies more than just an analogue representation of difference equations. He was convinced that the use of networks derives from the given data, (and from the given partial differential equations), *additional information* about the k-dimensional surface that was not contained in the second-hand difference equations; and which can be put to use, for instance, in the tearing apart and the piecewise solution of large networks. He conjectured that a network examined also the *neighborhood* of each given point (vertex) with the aid of currents and voltages residing in the added branches of the network. Following up this belief, the author eventually has generalized his one-dimensional physical models by a sequence of multidimensional networks, in order to derive still more information from the given data, or equations, with the aid of still more detailed neighborhoods.

The subject matter of the present paper utilizes practically the last step only in the generalization processes that discretize multidimensional fields by neighborhoods, rather than by mere points only. That step is a self-organizing "automaton." Thus it becomes necessary to give a bird's eye view of the various steps in the generalization process in order to appreciate the results of the last step, the oscillatory polyhedron or automaton.

Electric Circuit Models and Their Generalization

It is well known even in nonengineering circles, that the analytical solution to a *linear* field problem, say to the flow of heat within an odd-shaped region, may be approximated in a discrete manner by erecting between the assumed netpoints an irregular resistance network, and allowing the electrical currents in the added branches to imitate the flow and distribution of heat. (However, many mathematicians do not admit that the introduction of physics adds new information to the mathematical problem, since both approaches lead to exactly the same answer.) The next step in the functional generalization of such an analogue model consists of the use of inductors and capacitors, as well as ideal transformers, in addition to resistors. This step enables the simultaneous representation of several partial differential equations. (See Fig. 1.) A further generalization in the same direction undertakes the discretization of *nonlinear* field problems by means of *moving* networks, or moving currents measured along stationary networks, etc; such as arise in the study of rotating electrical machinery used in industry. This last generalization is equivalent to introducing two *spatially orthogonal*, one-dimensional networks and immersing them into a two-dimensional magnetic field traversed by

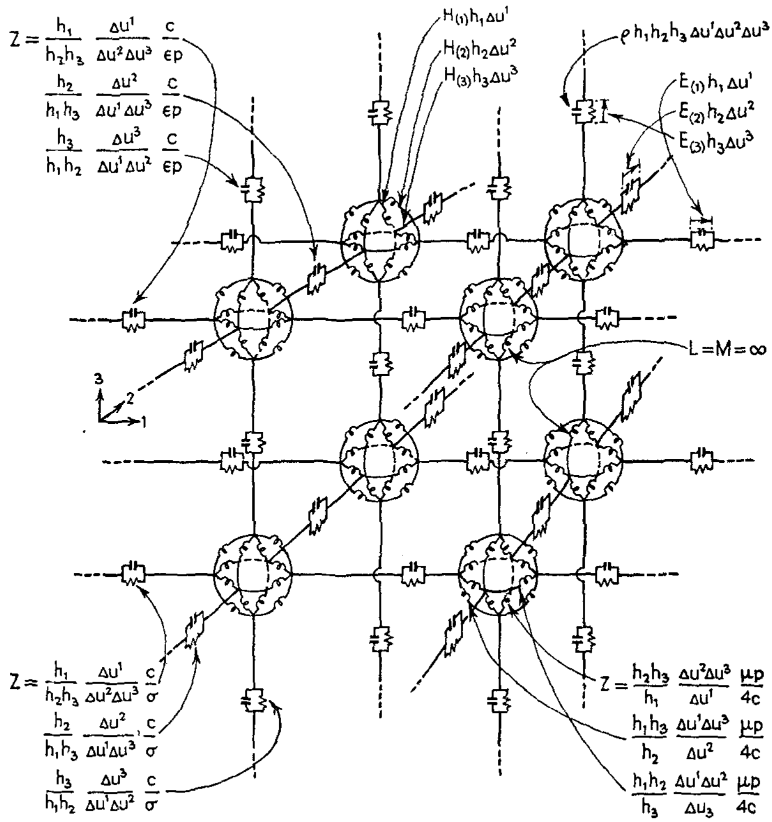


FIG. 1. Linear electric-network model of the field equations of Maxwell (orthogonal, curvilinear reference frame).

conduction currents. The networks act as nonholonomic reference frames and permit the analysis of the nonlinear field as a circuit structure by the addition of mechanical parameters (speed and torque) to the electrical parameters (current and voltage).

A more radical generalization along an entirely different direction consists of the enlargement of the one-dimensional (linear) stationary or moving network by polyhedral (nonlinear) networks (Fig. 2) having straight or curved multidimensional elements [2-4]. But now it becomes necessary to propagate across the polyhedron a sequence of electromagnetic waves, in place of conventional currents. The next logical and almost inevitable generalization is to immerse the polyhedron (and its dual polyhedron) into an underlying stationary or moving magneto-hydrodynamic plasma that fills a k-dimensional

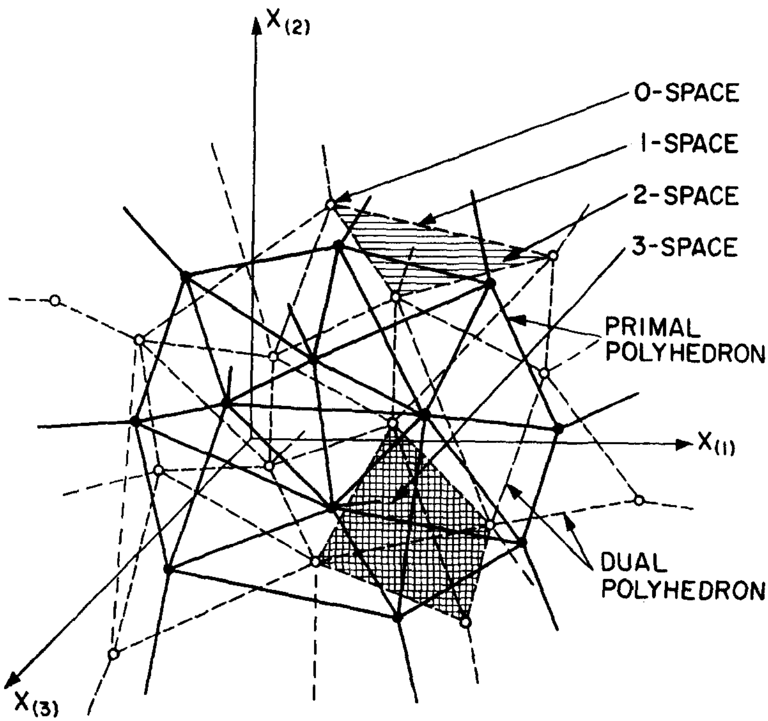


FIG. 2. Primal and dual polyhedra.

region. In addition to making available a hierarchy of additional electrical and mechanical parameters, the new kind of model combines the hitherto separate types of continuous and discrete approximations into one engineering tool. The resulting continuous plus discrete (crystallized) structure consists of a maze of multidimensional transmission networks (nonholonomic reference frames) connecting together a host of k-dimensional "generalized" rotating electrical machines, each having fluid and gaseous, electrostatic and electromagnetic, etc, sources, as well as conductors and dielectrics distributed continuously or discretely throughout a k-dimensional space.

Self-Organizing "Automata"

From the author's point of view, the advantage of the use of networks over netpoints is that each step in the above series of generalizations is introducing a large number of intrinsic *adaptive* parameters and feedbacks into the physical model. *The adaptive parameters are the carriers of all derived informations about the "neighborhoods" of the given points.* In the presence of extra parameters and thus of extra informations, more accurate and more versatile studies about the over-all problem become feasible.

A most fertile potentiality is opened up by the fact that the last-mentioned crystallized field structure may assume an *oscillatory* state under a great variety of boundary and environmental conditions. Hence it can be considered as a *self-organizing "automaton," possessing a great variety of intrinsic self-adaptive parameters and feedbacks.* Another potentiality is the fact that the automaton can also be given a deterministic or a probabilistic interpretation. The author expects to use the probabilistic automaton and its further generalizations for the solution of such cognitive processes as pattern recognition.

If the mechanical parameters (speed and torque) in the model are absent, the resulting stationary plasma and electromagnetic field can still oscillate in unison, and adjust themselves *automatically* to changes in their environment. This elementary nonmechanical automaton will be called an "oscillatory polyhedron." Its ability to curve-fit *any number* of arbitrarily picked multidimensional functions simultaneously is the subject matter of the present paper. It will be seen from the numerical examples to be shown, that the self-organizing (oscillatory) property of the structure is retained even in the absence of the motion of the underlying electric and magnetic charges and currents.

The automata mentioned above utilize only the minimum number of parameters and characteristics of a magnetohydrodynamic plasma. Actually it is possible to introduce an almost limitless sequence of thermodynamic and other parameters into the plasma, and thus into the automaton, in order to increase the versatility of the latter in solving complex "systems" problem.

Regression Analysis Versus Tearing Networks

Regression Coefficients as Liberated "Constraint" Variables

Although the introduction of higher order divided differences is part and parcel of the Calculus of Finite Differences, the curve-fitting will be performed not with Newton's interpolation formula, but with a generalized regression method employing a least-square criterion. The regression procedure will curve-fit not only a given function, but also its higher-order divided differences, all simultaneously. (It will be due to the oscillatory property of the polyhedron, and not to the generalization of the regression method, that the model will fit not one but any number of functions plus their divided differences.)

For several years the author has been developing a procedure—the method of tearing—to solve large-scale physical systems (linear, nonlinear, or oscillatory systems), in a piecewise manner [5]. He tore apart a physical system (or its network model) into n disjointed, isolated networks, and constructed a hypothetical "intersection" network out of the fragments of the torn n subdivisions. (The intersection network had the same eigenvalues as the original untorn system.) The author solved the $n + 1$ smaller networks separately, then interconnected the partial solutions into the solution of the resultant system. The resultant solution of the $n + 1$ smaller systems was as exact as if the original system had not been torn apart into n component systems, but solved in one piece.

It came as a pleasant surprise to discover that the philosophy underlying the theory of regression used in curve-fitting follows closely some of the steps that arise in the method of tearing. Both methods liberate hitherto hidden internal variables (regression coefficients residing within a hypothetical "intersection" system) that can throw *new light* upon the internal mechanism of the unknown overall system under study. Thus a general outline of the generalized curve-fitting procedure of this paper can be presented—ahead of the numerical examples—by comparing regression analysis with the method of tearing.

It may be mentioned that many people confuse the method of tearing networks (physical systems) with the method of partitioning matrices (set of equations) because of the similarity of their formulae. Such blind comparisons miss, however, the underlying basic differences. Tearing a network is like cutting open a patient; whereas partitioning a matrix (by any trick) is only like fluoroscoping the patient. A physical system (or its physical model) always contains more information than its mathematical representation, which uses only the minimum number of necessary equations.

Given Experimental Points

Let the numerical results of n experiments with k independent variables be expressed as an $n \times k$ data-matrix X. Of course, it is assumed that the selection of data satisfies all the requirements of *small sample theory*, whatever they may be, for the particular problem under consideration.

The matrix of independent variables X can be plotted as n points in a k-dimensional Euclidean space. The vector of dependent variables y represents points on the k-dimensional surface that has to be fitted by a model.

Developing the "Neighborhood" of Points

The immediate problem is to utilize the *neighborhoods* of the given n points for gathering more information about the unknown surface. For that purpose the author interconnects *neighboring* points (0-simplexes) with a network of lines, (branches or 1-simplexes), *neighboring* lines with a network of planes (triangles or 2-simplexes), and so on, until a network of k-dimensional Euclidean hyperplanes (k-simplexes) covers the entire k-dimensional space spanned by the n vertices (Fig. 2). The square of a q-dimensional volume element is called the "impedance" z of that element. Mutual impedances may also be introduced, representing products of direction cosines. A second, so-called "dual" polyhedron is also constructed, whose q-simplexes are orthogonal to those of the original, or "primal" polyhedron.

The existence of the polyhedron is equivalent to an enlargement of the original $n \times k$ data matrix X (k independent variables) by an equal number of rows and columns. Each added q-network contributes to X as many rows and columns as there are q-simplexes in the q-network.

The vector y of the dependent *variables* (the surface to be fitted) is also lengthened by propagating an electromagnetic wave across the torn-apart polyhedron. Hence the q-simplexes in a q-network must be so represented individually, and so interconnected with each other (and with the $q + 1$ or $q - 1$ simplexes of neighboring networks) that an electromagnetic wave with four types of parameters (e, b, h, d) can propagate across the network.

Satisfying Eight Maxwell's Equations Simultaneously

The linear network model that satisfies, in a discrete manner, the four field equations and the four constitutive equations of Maxwell, all simultaneously, is shown in Fig. 1, and is derived in [6]. The model contains resistors, inductors, and capacitors, as well as two types of ideal transformers. (The latter are necessary in order to take care of the *arbitrary* values of the variables.) Originally the author thought that for each higher dimension in the polyhedron a similar network would have to be constructed.

In attempting to satisfy Stokes' theorem between two different-dimensional networks, the author eventually discovered the fact (well known to geometers) that *even-dimensional spaces behave differently from odd-dimensional spaces*, and thereby, in a polyhedron, *two complete different-dimensional networks are necessary to generate one complete electromagnetic wave*. Thus a k-dimensional polyhedron includes a sequence of only $k/2$ full waves. (This is the reason why the simplest possible example that can be worked out with the aid of a complete electromagnetic wave must be two-dimensional and not one-dimensional. A planar network must also be used.)

All even-dimensional q-networks were thus constructed of magnetic material (inductors and resistances), and all odd-dimensional q-networks of dielectric material (capacitors and conductances). In order to introduce ideal transformers, it was necessary to assume a dual polyhedron also, in which the physical role of even and odd dimensions are interchanged. The two polyhedra also are interconnected both conductively and inductively. It may be stated that *each full electromagnetic wave (with e, h, b and d components) still occupies the configuration of Fig. 1*, except that now the configuration extends over both a q-dimensional and a $q+1$-dimensional network.

The Method of Tearing

Before energizing the polyhedra, two preliminary steps implied by the method of tearing must be performed:

1. Each polyhedron is torn into isolated q-networks (separate linear-, triangular-, tetrahedral-, etc. networks), then their impedance matrices are established and inverted.

2. Out of the fragments of the inverted and isolated q-networks two additional polyhedra, the so-called "intersection" polyhedra are constructed (both primal and dual).

These component networks are separately energized.

Regression Analysis

The determination of the model-errors in curve-fitting by regression consists of the same steps that arise in solving *piecewise* the linear equations of state $I = YE$ of a network by the method of tearing. (In the linear network, the known quantities are the impressed "open-path" currents I; whereas the unknowns are the potential differences E appearing across the same open paths.) If the terminology of linear networks is used for a simplified presentation of the polyhedral network, the steps in the piecewise solution of a polyhedron are as follows:

1. The given dependent variables y are considered as known currents I impressed upon the vertices of the untorn polyhedron. (The vertices belong to the 0-network and not to the linear-, or 1-network.)

2. The polyhedron is torn into isolated q-dimensional networks (subdivisions) and the open-circuit voltages E' appearing upon each q-network are calculated. These voltages are considered to represent the "divided differences" (approximations to directional derivatives), that also have to be fitted by the polyhedron, simultaneously with the function. The vector of open-circuit voltages E' forms the extended y vector.

3. The open-circuit voltages are impressed as $e = X_t y$ upon the intersection polyhedron whose impedance matrix is $X_t X = S$.

4. The intersection-network impedance matrix is inverted as

$$(X_t X)^{-1} = S^{-1}.$$

5. The resultant currents i are the regression coefficients $\hat{\beta}$ of the least-square estimation. $\hat{\beta} = (X_t X)^{-1} X_t y$ or $i = (X_t X)^{-1} e$.

6. The additional currents I' that appear upon the isolated subdivisions (because of the existence of the intersection network) are $\hat{y} = X\hat{\beta}$. The estimated function and all its estimated divided differences $\hat{y}$ are thereby found.

7. The difference between the given and estimated "function plus divided differences" $E' - I' = y - \hat{y} = \epsilon$ is the error of the polyhedral model in fitting the given function and all its divided differences simultaneously.

8. The unknown potentials E appearing upon the vertices of the untorn polyhedron are not utilized in the present study.

Thus the physical tearing apart of the original polyhedron is an absolutely necessary procedure, since it is the hypothetical "intersection" polyhedron that acts as an estimating model. The tearing liberates new, otherwise hidden constraint forces that play the role of the unknown structural concepts sought.

Nature of Polyhedral Curve-Fitting

The curve-fitting process does not involve any reduction in the original degrees of freedom, since X is enlarged by *equal* number of rows and columns. That is, the improvement in curve-fitting is now based upon the utilization of the divided differences of the function (rows added to X) and not upon the reduction of the degrees of freedom of the data-matrix by adding only columns to X.

It should be remarked that the data-matrix, hence the conventional regression procedure, also enters in the form of a (-1)-network attached to the polyhedron. As a result, if so desired, it is also allowable to guess an arbitrary functional form that fits the dependent variable. (The guess changes the original data-matrix and thus it may reduce the degrees of freedom.) An absence of guess is equivalent to a hyperplane guess. *The role of the polyhedron is to improve the fit given by the conventional regression model by introducing more regression coefficients, but without reducing the existing degrees of freedom.* Thus, when only a limited number of points are available, a k-dimensional polyhedron can take the place of a polynomial with k variables and high degree.

The polyhedral model is still linear in the regression coefficients β. It is, however, nonlinear in the independent variables x, since the volume elements are involved polynomials of x.

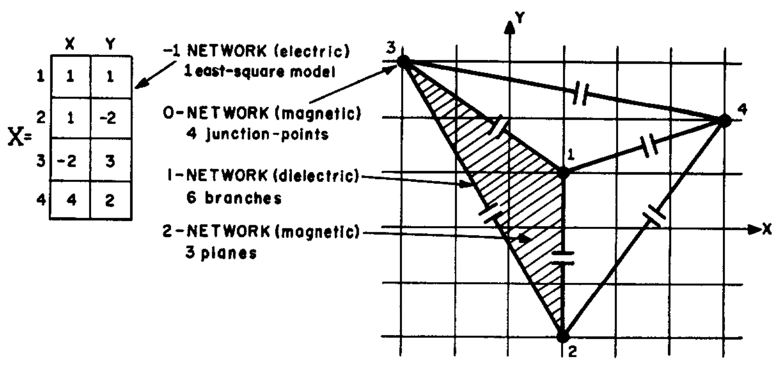
Numerical Examples

Fitting with a Nonoscillatory Polyhedron

The original aim of the author was to show that a polyhedron can always improve the fit of a k-dimensional function made by conventional regression theory, *without reducing the available number of degrees of freedom.* The appearance of fitted divided differences was only an added byproduct.

Since it was found that different interconnections gave different fitting for the same function, also for different functions, it became the goal of a later research to discover whether one and the same model would fit, in a satisfactory manner, not one but any number of arbitrary functions simultaneously, without changing the model. Thus the examples to follow in this section refer to trials with a *nonoscillatory* polyhedron only, whose behavior can be explained without introducing a magneto-hydrodynamic plasma and thus an oscillatory, self-organizing polyhedron.

(A) *Given data-matrix*



(B) *Given trial functions*

(1) $z_1 = 3x + 4y + 2$
(2) $z_2 = 5xy + 3x + 4y + 2$
(3) $z_3 = x^2 + 2y^2 - 3xy + 5x - 3y + 6$
(4) $z_4 = x^3 + 2y^2x - 2x^2 - 2xy + 4y + 8$
(5) $z_5 = 10[1 + \log(x + 4) - 1/y]$
(6) $z_6 = 10[2 + e^{x/10} - y \log(x + 10)]$

Fig. 3. Three-dimensional curve-fitting with a polyhedron
(Simultaneous fitting of a function plus all its divided differences.)

Derived Characteristics of a Function

Four points were assumed as discrete independent variables upon a two-dimensional plane ($n = 4$, $k = 2$), shown in Fig. 3. The polyhedral reference frame attached to the given points contains six branches and three triangles (planes), along which six first-order and three second-order divided differences can be defined. The assumed model can consider thus $4 + 6 + 3 = 13$ characteristics of the given function (or rather, of the discrete data) for curve-fitting or other purposes.

The $13 - 4 = 9$ derived characteristics represent a minimum. Their number can be increased by assuming still higher-order divided differences (just as in the Calculus of Finite Differences). By subdividing (triangulating) the polyhedron, any desired number of further characteristics of a given function can be derived. Moreover with each divided difference the components of a skewsymmetric tensor of rank q are also associated, representing still additional characteristics, namely projections of the directional derivative upon various coordinate hyperplanes (Grassman coordinates).

As the assumed system has one residual degree of freedom ($n - k - 1 = 1$), the 12×12 moment-matrix can supply 12 regression coefficients to estimate the characteristics of a particular function. In working out numerical examples six quite different functions were assumed to generate regression coefficients. It was found by trial that each nonoscillatory polyhedron could fit satisfactorily only *one* of the functions and its divided difference; at most a restricted class of functions only. It was found that in a given polyhedron different reference frames give different fit. The results of two different *planar* reference frames will now be shown, assuming the reference frame of the linear network to be unchanged.

The six assumed functions and their derived first and second divided differences are listed in the left-hand columns of Tables I to VI.

Estimation by a Hyperplane

In fitting only *one* dependent variable $z = f(x, y)$, it is possible to guess any arbitrary function that may fit z by least square. In the present example a large number of dependent variables exist, hence no single function could be guessed at. Thus for each of the six z functions a *hyperplane* fit $z = a + bx + cy$ is being applied automatically by the -1 network of the polyhedral model.

Considering a hyperplane model as a first guess to fit each of six functions, the conventional least-square estimates are given in the first right-hand columns of Tables I to VI. The least-square estimation does not utilize any divided differences.

TABLE I

$z = 3x + 4y + 2$

	Three functions to be fitted	Least-square	Least-square + linear	Least-square + linear + planar	Least-square + linear + planar	Least-square + linear + planar
Function	9.0000000	9.0000002	9.0796617	8.8787335	8.9632293	8.9999997
	— 3.00000000	— 2.9999995	— 3.0265549	— 2.9595791	— 2.9877437	— 2.9999993
	8.0000000	8.0000000	7.9734455	8.0404218	8.0122575	8.0000011
	22.000000	22.000000	21.973445	22.040421	22.012257	22.000001
First div. diffs.	4.0000000		5.9057718	4.0538963	4.0649511	3.9999975
	1.8864844		— 1.8864846	1.8795529	2.0501086	1.8864869
	5.0000000		5.0000005	4.9757475	4.8752506	4.9999983
	— 0.27735011		— 6.4341227	— 0.32219431	— 0.410672249	— 0.27734973
	4.1109605		6.5376262	4.0598305	4.0416890	4.1109617
	2.3015856		— 2.3015861	2.3148763	2.4084983	2.3015888
Second div. diffs.	— 4.5000000			— 4.4910178	— 4.3879722	— 4.5000036
	— 1.4558823			— 1.5007952	— 1.5821587	— 1.4558866
	— 1.7027027			— 1.7206683	— 1.7817965	— 1.7027075
				Two nonoscillatory polyhedra		Oscillatory polyhedron

TABLE II

$z = 5xy + 3x + 4y + 2$

	Three functions to be fitted	Least-square	Least-square + linear	Least-square + linear + planar	Least-square + linear + planar	Least-square + linear + planar
Function	14.000000	10.250000	11.990840	11.562751	13.260975	14.000001
	— 13.000000	— 11.749999	— 12.339281	— 12.187586	— 12.753653	— 13.000004
	— 22.000000	— 20.750002	— 21.330283	— 21.187583	— 21.753650	— 22.000000
	62.000000	63.250002	62.669720	62.812415	62.246347	62.000001
First div. diffs.	9.0000000		8.3047299	10.083224	9.7053922	9.0000045
	— 1.5434872		1.5434872	— 1.6828137	— 1.7534652	— 1.5434868
	15.000000		15.000005	14.512556	13.498241	15.000001
	— 9.9846040		— 29.007562	— 10.885893	— 10.500833	— 9.9845986
	15.178931		28.690252	14.151301	15.051589	15.178937
	13.809513		— 13.809516	14.076639	13.525182	13.809518
Second div. diffs.	— 13.500000			— 13.319460	— 13.948425	— 13.500002
	1.1911765			.28848957	.45322337	1.1911775
	— 10.216216			— 10.577281	— 10.905862	— 10.216220
				Two nonoscillatory polyhedra		Oscillatory polyhedron

TABLE III

$$z = x^2 + 2y^2 - 3xy + 5x - 3y + 6$$

	Three functions to be fitted	Least-square	Least-square + linear	Least-square + linear + planar	Least-square + linear + planar	Least-square + linear + planar
Function	8.0000000	21.750000	14.084910	13.529867	9.6767745	7.9999970
	32.000000	27.416667	29.971697	30.156712	31.441073	32.000001
	27.000000	22.416667	24.971695	25.156712	26.441073	27.000001
	20.000000	15.416667	17.971694	18.156712	19.441073	20.000000
First div. diffs.	— 8.0000000		— 11.586759	— 10.457718	— 8.5617813	— 7.9999956
	— 0.85749290		0.85749275	— 0.54137159	— 0.15564890	— 0.85749047
	— 2.4000000		— 2.3999995	— 1.2940275	— 0.72377002	— 2.4000035
	5.2696521		7.8236599	7.3145987	6.3015391	5.2696505
	3.7947328		5.1638820	6.1263263	5.0562755	3.7947320
	— 1.1507928		1.1507931	— 1.7568613	— 2.3764763	— 1.1507943
Second div. diffs.	2.1600000			1.7503805	3.3513939	2.1600005
	0.66176469			2.7098610	3.7201204	0.66175650
	0.85135136			1.6705914	0.85810646	0.85135590
				Two nonoscillatory polyhedra		Oscillatory polyhedron

TABLE IV

$$z = x^3 + 2y^2x - 2x^2 - 2xy + 4y + 8$$

	Three functions to be fitted	Least-square	Least-square + linear	Least-square + linear + planar	Least-square + linear + planar	Least-square + linear + planar
Function	11.000000	16.500000	12.867435	12.176993	11.356887	10.999996
	11.000000	9.1666670	10.377522	10.607675	10.881035	10.999999
	— 20.000000	— 21.833335	— 20.622481	— 20.392334	— 20.118969	— 19.999999
	64.000000	62.166669	63.377522	63.607662	63.881024	64.000000
First div. diffs.	0.0000000		— 6.3005580	— 0.52311256	0.56960430	0.00000180
	— 5.3164560		5.3164556	— 5.2491730	— 5.1553025	— 5.3164533
	10.600000		10.600001	10.835401	9.8081517	10.599998
	— 8.5978534		— 30.853252	— 8.1625995	— 8.3855028	— 8.5978524
	16.760070		31.905905	17.256334	17.748653	16.760080
	13.809513		— 13.809516	13.680515	12.163551	13.809517
Second div. diffs.	— 9.5400000			— 9.6271853	— 9.2773363	— 9.5400275
	4.1029412			4.5388641	5.7785725	4.1029362
	— 10.216216			— 10.041846	— 11.248539	— 10.216220
				Two nonoscillatory polyhedra		Oscillatory polyhedron

Linear-Network Estimation

Of course, the entire polyhedron can be immediately attached to the regression (-1 dimensional) network and the final estimate calculated outright. As a matter of fact, even the hyperplane fit need not be calculated as an intermediary step. Nevertheless, in order to discover the relative effectiveness of the sequence of q-networks, next only a 0-network plus a 1-network were attached to the (-1)-network. The results of the second step are shown in the second right-hand columns of Tables I to VI.

The estimate of the fundamental functions became improved in all cases by the addition of a linear network. An approximate estimate of the first divided differences was also found.

Planar-Network Estimations

When three planes were added to the linear network, the polyhedral model attached to the least-square network represented one *full* electromagnetic wave. (The polyhedral model comprised part of a 0- and a 2-network, plus an entire 1-network.) Two different reference frames were assumed on the planar network.

The estimated functions are given in the third and fourth right-hand columns of Tables I-VI. The second planar reference frame gave much better fit than the first frame, with several of the functions.

Goals Accomplished

The cited numerical examples of Tables I-VI show that a polyhedron attached to a least-square model accomplishes the expected goals:

1. It improves *always* the estimate of the conventional least-square model without decreasing its degrees of freedom.

2. It gives simultaneously satisfactory estimates for the higher-order divided differences also.

In analogy to a conventional regression model, any guessed-at reference frame upon the polyhedron can be improved only by further guessing, or by a systematic study (for instance, by introducing still higher-order divided differences). Thus each class of dependent variables requires, in general, a different polyhedral model, just as in conventional curve-fitting by regression.

TABLE V

$$z = 10[1 + \log(x + 4) - 1/y]$$

	Three functions to be fitted	Least-square	Least-square + linear	Least-square + linear + planar	Least-square + linear + planar	Least-square + linear + planar
Function	6.9897000	13.171826	9.5941080	9.3645221	7.7098022	6.9896970
	21.989700	19.928993	21.121563	21.198096	21.749675	21.989704
	9.6770000	7.6162930	8.8088612	8.8853929	9.4369703	9.6770038
	14.030900	11.970192	13.162762	13.239290	13.790879	14.030902
First div. diffs.	− 5.0000000		− 8.3303463	− 6.0554783	− 4.7719472	− 4.9999998
	− 2.1116106		2.1116106	− 1.9758515	− 1.2343125	− 2.1116095
	− 1.5917600		− 1.5917597	− 1.1167957	− 1.6540805	− 1.5917634
	0.74532295		− 1.7515334	1.6235324	0.83237203	0.74532097
	2.2266227		4.0689686	3.2279351	2.9507748	2.2266213
	0.71577668		− 0.71577673	0.45549796	− 0.16140631	0.71577594
Second div. diffs.	1.4325840			1.2566711	2.3886725	1.4325833
	1.6296221			2.5091867	3.2025729	1.6296211
	− 0.52952837			− 0.17770240	− 1.2305923	− 0.52953138
				Two nonoscillatory polyhedra		Oscillatory polyhedron

TABLE VI

$$z = 10[2 + e^{x/10} - y \log (x + 10)]$$

	Three functions to be fitted	Least-square	Least-square + linear	Least-square + linear + planar	Least-square + linear + planar	Least-square + linear + planar
Function	20.638100	21.401976	20.392579	20.674573	20.649166	20.638098
	51.879800	51.625176	51.961641	51.867649	51.876114	51.879799
	1.0946000	0.83997600	1.1764357	1.082441	1.0909113	1.0946006
	11.995400	11.740776	12.077236	11.983245	11.991711	11.995401
First div. diffs.	— 10.413900		— 19.266877	— 10.430116	— 10.733432	— 10.413894
	— 8.7095897		8.7095898	— 8.7075051	— 9.3419035	— 8.7095875
	— 7.9768800		— 7.9768800	— 7.9695840	— 7.4394390	— 7.9768850
	— 5.4203919		— 18.406788	— 5.4069027	— 5.0197307	— 5.4203946
	— 2.7330614		— 0.97326632	— 2.7176835	— 2.7122249	— 2.7330566
	1.7920803		— 1.7920805	1.7880821	1.7599211	1.7920777
Second div. diffs.	7.1791920			7.1764900	6.6954955	7.1791938
	6.7215708			6.7350803	6.7595548	6.7215713
	— 1.3257730			— 1.3203688	— 0.85198171	— 1.3257760
				Two nonoscillatory polyhedra		Oscillatory polyhedron

Almost-Oscillatory Polyhedron

Search for a Universal Model

However, it was the eventual goal of the author that the same model, *without adjustment* and without being of undue size, should fit:

(1) *Any number* of arbitrary dependent functions, all simultaneously.

(2) Functions with *unknown components* or unknown nonlinear parameters.

(3) Also it should fit simultaneously all higher-order divided differences, integrals, as well as functions of the latter (if given).

To satisfy such a large order, it occurred eventually to the author that the polyhedron itself (without the regression-network or — 1 network) should be made oscillatory, or self-excited. Such a stage could be expected to be reached by enlarging the purely electromagnetic wave (b, d, e, h) with electric and magnetic charges and currents (ρ^e, J^e, ρ_m, J_m). This next step in the generalization was suggested by the unused portions in the configuration of Fig. 1, which still had just enough room left to incorporate the added electrical source variables (without the mechanical variables). The complete Fig. 1 appears to be a well-balanced structure, that might oscillate in the proper polyhedral surroundings, if stretched into two dimensions. It was hoped that, after adding the nonoscillatory least-square network, the resultant polyhedron (an almost oscillatory structure) will not be unduly influenced by the nature of the functions that it is supposed to estimate.

Estimation with an Oscillatory Polyhedron

The last columns in Tables I-VI show the improved estimates given by an oscillatory polyhedron for the same six functions and differences that are given in Fig. 3. The previous linear network configuration was retained also in the oscillatory model, and only the planar network configuration was revamped. The estimates came out even better than the author expected. It is, however, emphasized that *no exact fit is possible*, since the polyhedral model still has one residual degree of freedom, the same that the least square model has.

The good estimates of six widely different functions suggest that *no guessing nor adjusting of the oscillatory polyhedral model is required when the dependent variables differ widely*. Because of this elastic, self-adjusting feature of the almost-oscillatory polyhedron, the dependent variables that are to be fitted can thus be arbitrary functions, or multivariate functions; they can have unknown components, or can be functions of *unknown* parameters.

The almost-oscillatory polyhedron is expected to be a universal model for estimation in a large class of problems.

Interpolation

Thirteen characteristics of the given functions and an already inverted 12×12 moment-matrix are now available for further studies with the six functions, such as interpolation, smoothing, generalized harmonic analysis, etc. As a matter of fact, more than thirteen characteristics of each function are known, since, for instance, with each set of divided differences a skew-symmetric tensor of rank q (containing the Grassmann coordinates) is also associated.

A whole gamut of crude to refined methods can now be employed for interpolation. The already known twelve regression-coefficients of each function may also be used either unchanged or in an altered form, as additional interpolated points are assumed either singly or in groups.

It is emphasized that the polyhedron interpolates not only points, but at the same time also their entire neighborhood by means of interpolated branches, triangles, tetrahedra, etc. Thus *the oscillatory polyhedron calculates not only the unknown functions at the interpolated points, but simultaneously also all their higher-order divided differences* in the directions defined by the interpolated q-networks. It is also emphasized that the multidimensional interpolation with a polyhedron is subject to the same analytical limitations as the one-dimensional interpolation of the Calculus of Finite Differences. The polyhedral interpolation should not be expected to accomplish theoretical feats that the one-dimensional Calculus of Finite Differences cannot do.

Physical Interpretation

The propagation of electromagnetic waves in *multidimensional* networks offers many unusual and unexpected features that lie outside the experience of the author. Some aspects of the physics are still puzzling, but the numerical results speak for themselves. Hence the physical interpretation given should be considered as only tentative, until more numerical data are gathered about the nature of propagation of a sequence of waves across a sequence of higher-dimensional networks.

It should also be pointed out that the given numerical examples (fitting six arbitrary functions right at the given points) were devised merely as a convenient *check* on the correctness and oscillatory nature of the polyhedral model, and not as a field of possible application. After all, there exists an infinite number of data-matrices that can produce exact fits at the four given points for all six functions, (even for their divided differences), without any error. (Of course, no degrees of freedom would be left over as occurs in these examples.) It is the quality and quantity of errors that determine the value of a model and not the absence of all errors. The polyhedral model is the depository of several other types of errors also, that have not been discussed in these pages.

Minimum Stored Energy

Following up the physical analogy between regression theory and its network analogue, it can be shown that the total power input into the network represents the error-square. Hence *the model error is a minimum, if the total stored electromagnetic energy in the resultant polyhedron is also a minimum.* If the intersection network "explains" totally, or fits perfectly, the given function and all its divided differences, then the stored energy in the interconnected polyhedron is zero.

In a linear network the stored energy is zero only if the network is oscillatory. Thus before excitation both the resultant polyhedron and the intersection polyhedron must be made oscillatory. That is, the reference frames in the individual q-networks and in the intersection polyhedron must be so assumed in unison, that the resultant over-all waves propagating within the interconnected polyhedron should be self-sustaining in space and time. This last step of finding a set of open-circuit and short-circuit reference frames that have minimum stored energy when combined, is basically a nonlinear programming problem. Because of the simplicity of the problem, however, the best reference frame may be arrived at by physical reasoning, without a nonlinear programming process. (This physical reasoning has not been reduced as yet to a routine computer procedure.)

The regression network itself (-1 network) cannot be oscillatory when it has one or more degrees of freedom; thereby the polyhedron as a whole can only *approach* an oscillatory stage, but can never reach it. Thus the errors may approach zero much faster if the data-matrix itself approaches a square.

Crystals as Self-Organizing Automata

It is emphasized, however, that the oscillations in a polyhedron represent a far more advanced type of physical phenomenon than do the oscillations in a conventional electrical network. In the "self-organizing automaton" the oscillations are self-adaptive, that is, they can adjust their modes and frequencies automatically to varying conditions. It is an inherent property of all electrified polyhedral networks that the manner of interconnection of the underlying *material* p-simplexes is not permanent, but varies with the nature of the superimposed *electrical* waves. (This is the motivation behind the author's concept of "tearing and interconnecting" electrified networks [8, 9].) On the other hand, the oscillations in a conventional "oscillatory network" are not self-adaptive, but fixed in mode of oscillation and frequency. Fortunately nature offers many examples of self-adaptive electrified structures.

The atoms of a polyatomic molecule within a crystal may be assumed to represent the vertices of a polyhedron. When a crystal is excited by x-rays, the resulting dipole waves of the oscillating atoms and their accompanying diffracted electromagnetic waves form a self-sustaining dynamical system. The multidimensional waves in the oscillatory polyhedron of this paper are surprisingly complete analogues of such resonant crystal-waves. (However, the curve-fitting examples utilize ray-optics rather than wave-optics.)

Thus it lies within the realm of possibility that crystals will be employed eventually for the physical realization of multidimensional polyhedral networks energized with electromagnetic, magnetohydrodynamic, and still more advanced types of waves. It appears that these advanced types of electromagnetic waves can offer a greater variety of parameters (and thus more subtle multivalued logics) for the construction of crystal computers than switching-circuit signals can offer in conventional computers. The existence of electromagnetic-wave propagation along a neural network has already been long established. Thus a crystal may also act as a model for some classes of neural phenomena.

Summary

Summary of Nonoscillatory Polyhedron

1. The Calculus of Finite Differences deals with functions, as well as with their divided differences of higher order, but in one dimension only. Two or three dimensions are treated only as combinations of, or iterations upon, one-dimensional projections. (For a detailed discussion of this important point by Salzer, see [7].)

2. The regression theory deals with functions in any number of dimensions, but ignores their divided differences.

3. The animated polyhedral approach deals with functions, as well as with their higher order divided differences, all simultaneously, in any number of dimensions.

The polyhedral approach can thus deal adequately with problems in which a *limited* amount of data is associated with a *large* number of independent variables. However, a nonoscillatory polyhedron, as well as a conventional nonlinear regression model, can deal, without adjustment, with only a small class of functions at a time. In the presence of additional arbitrary dependent variables additional regression models need to be guessed at; and in the polyhedron a different model or a different reference frame needs to be assumed.

Summary of Almost-Oscillatory Polyhedron

The self-organizing, oscillatory capability of a polyhedral model (automaton) injects some unusual features into the problem of estimation. For instance:

1. A single oscillatory polyhedron can estimate not only *one* function and its associated family of divided differences, but *any number* of arbitrarily selected functions and their divided differences, all simultaneously.

2. The oscillatory polyhedron can estimate satisfactorily even a set of random numbers. (The significance of this is not yet understood.) Hence if the multivariate dependent variables contain *unknown* components or unknown linear parameters, the model can still estimate and solve them.

3. If the unknown parameters of the multivariate dependent variables are nonlinear, the estimating polyhedron establishes a set of nonlinear algebraic relations between the parameters. However, a second oscillatory polyhedron is necessary to solve the nonlinear algebraic equations.

Acknowledgment

The author wishes to acknowledge his indebtedness to James L. Dolby for initiating and directing the present problem with numerous suggestions. The author is also indebted to Dr. W. J. Merrill for comments and criticisms, and to E. G. Bianco for his interest.

References

1. Hodge, M. V. D. "The Theory and Application of Harmonic Integrals." Cambridge Univ. Press, 1952.
2. Kron, G. A generalization of the calculus of finite differences to nonuniformly spaced variables. *Trans. AIEE Pt. I.* **77**, 539-544 (1958).
3. Kron, G. Multidimensional space filters. *Matrix and Tensor Quart. (London).* **9**, No. 2 (1958).
4. Kron, G. Basic concepts of multidimensional space filters. *Trans. AIEE, Pt. I.* **78**, 554-561 (1959).
5. Kron, G. Diakoptics—The piecewise solution of large-scale systems. A serial of 20 chapters in the June 7, 1957 to Feb. 13, 1959 issues of *Elec. J. (London).*
6. Kron, G. Equivalent circuits of the field equations of Maxwell-1. *Proc. IRE.* **32**, 289-299 (1944).
7. Salzer, H. E. Some new divided difference algorithms for two variables. Proc. of Symposium, Mathematics Research Center, U.S. Army at the University of Wisconsin, Apr. 21-23, pp. 61-98 (1958).
8. Kron, G. Camouflaging electrical 1-networks as graphs. Scheduled to appear in the *Quart. Appl. Math.*
9. Kron, G. "Tensors for Circuits." Dover, New York, 1959.