Visibility Laboratory
University of California
Scripps Institution of Oceanography
San Diego 52, California

TEMPORAL METRIC SPACES IN RADIATIVE TRANSFER THEORY

I. TEMPORAL SEMIMETRICS

Rudolph W. Preisendorfer

1 January 1959
Index Number NS 714-100

Bureau of Ships
Contract NObs-72092

SIO REFERENCE 59 - 2

Approved:

Seibert Q. Duntley

Seibert Q. Duntley, Director
Visibility Laboratory

Roger Revelle

Roger Revelle, Director
Scripps Institution of Oceanography

Temporal Metric Spaces in Radiative Transfer Theory

I. Temporal Semimetrics

Rudolph W. Preisendorfer

Scripps Institution of Oceanography, University of California

La Jolla, California

INTRODUCTION

The ultimate purpose of this study is to lay the groundwork for a topological characterization of the general time-dependent transport processes common to the theory of radiative transfer and neutron transport phenomena. In this way the recent formulations of topological dynamical analysis[1] will be made available for the general study of these phenomena, and will result in a new set of theoretical tools for their respective theories. For the present, however, our goal is more modest. In the first part of the study we will be concerned with the introduction and illustration of the notion of the temporal semimetric on a general carrier space.[2] The subsequent four parts of the study will consider various extensions and applications of this concept to the general time-dependent multiple scattering problem in radiative transfer theory. Besides preparing the groundwork for the ultimate goal, the five parts of the present research program provide a set of interesting facts about time-dependent radiative transfer processes which may be studied for their intrinsic interest.

Before going into the initial details, we consider some of the current approaches to the time-dependent multiple-scattering problem.

Direct Approach

The time-dependent multiple scattering problem in radiative transfer theory may be explored from several points of view. The <u>direct approach</u> in its most general aspects starts with the general time-dependent form of equation of transfer[2]

$$\left(\frac{n^2}{v}\right) \frac{D[N/n^2]}{Dt} = -\alpha N + N_* + N_\eta \ ,$$

where n is the index of refraction function, v is the speed of light function, α is the volume attenuation function, N is the radiance function, N_η is the emission function, and N_* is the path function defined by

$$N_* = \int_{\Xi} \sigma N \, d\Omega \ ,$$

where σ is the volume scattering function. It is assumed that n (and hence v), α, σ, and N_η are given on some carrier space for each time in some given interval of time (t_0, t_1). N is assumed known at t_0 at all points of the carrier space. The main goal is to determine N on the carrier space for all times t' in the interval (t_0, t_1).

Invariant Imbedding Approach

An alternate approach to the time-dependent multiple scattering problem is the so-called invariant-imbedding approach[3] which assumes the existence of a certain set of generalized reflectance and transmittance operators on the space of radiance functions. These operators are shown to satisfy a certain set of functional equations which in principle may be solved to yield the desired solution on wide range of optical media. This approach is the time-dependent extension of the classical steady state procedures first explored by Ambarzumian,[4] and later developed in detail for certain simple contexts by Chandrasekhar.[5]

Disadvantages of the Above Approaches

Both the direct approach and the invariant imbedding approach require an extensive calculation program (either theoretical or numerical) to arrive at the required solutions before they yield even the simplest details of the time-varying radiation field. Thus while we possess two fundamentally powerful and exhaustive analytical approaches to the time-dependent multiple scattering problem, the information held within their formalisms is locked from view by formidable analytical barriers. For even the simplest of contexts, the most elementary properties of the solutions must await the <u>complete</u> solution before they are accessible to view.

Divide and Conquer Principle

One way to provide a remedy for this state of affairs within radiative transfer theory is to adopt the Divide and Conquer Principle. It analyzes a problem into a formal sequence of partial yet exactly formulated problems such that the "sum" of the solutions of the partial problems equals the solution of the whole problem. In particular the single infinitely difficult multiple scattering problem can be analyzed into an infinite number of easily solved partial problems. The accumulation of solutions to the easy problems can be compounded as long as the will and strength of the investigator holds out. Any ground gained as progress is made along the sequence of problems is ground irreversibly won. In many instances, the main features of the solution are discernable after only a few partial problems are solved.

n-ary Scattering Approach

An explicit example of the general Divide and Conquer Principle is the <u>n-ary scattering approach</u> in which the radiance associated with single-scattered radiant energy is precisely determinable once the initial radiance conditions are stipulated. From this, in turn, the radiance of secondary-scattered radiant energy is precisely determinable. By induction, once n-scattered radiance is known, $(n+1)$-scattered radiance is precisely determinable, where n is any integer. The required radiance function is the sum of the convergent infinite series whose nth term is the n-ary radiance function.

Some workers are of the opinion that such an approach is too elementary, too straight forward, or too laborious. For them, it apparently is an uncomfortable realization that such an infinitely complex problem has such a conceptually simple and completely determinate solution procedure as the n-ary approach. When pressed for a reason for their opinion they say that the n-ary approach negates the entire purpose of the integro-differential formulation of the problem[6] (the "direct" approach defined above). This is a good answer, provided one shares a comparable reverence for the integro-differential formulation.

However, anyone who has given more than cursory thought to the multiple scattering problem, and to the problems of mathematical physics in general should arrive at a less superficial judgment of the n-ary scattering approach. A more favorable judgment may be based on the following three observations: First, the n-ary scattering approach is a concrete physical realization of the Neumann iterative procedure for solving very general types operator equations.[7] This method is at once simple and far-reaching in its applicability. Not only does the multiple scattering problem of radiant energy fall under its extensive domain, but also every physical phenomenon represented by a functional equation (which, incidentally, just about exhausts the content of all classical and modern theoretical physics). Second, the advent and rapid development of modern large scale electronic computors has provided a natural vehicle for the numerical solution of the Neumann formulation of physical problems. The basic iterative nature of the Neumann formulation is natural grist for these electronic mills. Once one of the larger of these mills has been supplied with the appropriate program, the most formidable member of the species of multiple scattering problems will yield to tabulation in a matter of hours. Finally, the majority of workers who adopt the direct approach, whether they deign to admit it or not, clearly negate its inherent strength by first assuming extremely simple and physically meaningless assumptions, (the most objectionable of these is the isotropic-scattering assumption) and then handing over the emasculated problem to large scale computors for numerical solution. The point here is not that simplifications are to be avoided in theoretical studies, but rather that simplifications should never be used as unnecessary means to an end. If the solution of the multiple scattering problem is to eventually end up in numerical tabulations, it should then be done once and for all with the most exact and powerful approach currently possible, using realistic physical assumptions. The momentary cost to accomplish this would be very high, but in the long run, it would be the most economical and sensible step to take. If all the energy lavished so far on the currently extant isotropic tabulations in both radiative transfer and neutron transport theory is totalled, there would be enough energy to solve at least a dozen well-chosen sets of anisotropic, boundary-value problems — enough to supply both theoreticians and applied workers in both fields with examples of the true power of their mathematical models so that, perhaps for the first time in each theory, crucial tests can be made of their general validity.

The Present Approach

The present approach belongs to the class of approaches under the Divide and Conquer Principle. The approach leads to a minute and exact examination of a very small subdomain of the general multiple scattering problem. The basic idea behind the approach has the appealing simplicity of the $\cap$-ary scattering methodology, and is endowed with the same potentialities: by an iterated sequence of applications it is capable of a complete solution of its assigned share of the basic problem.

Specifically, the present approach will be used to ask and answer in complete detail the following five questions:

Given a general carrier space with a prescribed set of initial conditions:

(a) What analytic tool will allow the representation of the general radiation field in terms of the more common notions extant in classical (euclidean) carrier spaces? The answer is supplied by the concept of the temporal metric, and the initial stages of the answer will be given in the present paper.

(b) At any given instant, what <u>points</u> in the carrier space are actively contributing radiant flux to a <u>given</u> point in the space? The answer may be phrased in terms of the characteristic function. Details are to be found in the second paper of the series.

(c) What <u>regions</u> in the carrier space are actively contributing radiant flux to, and actively receiving flux from a <u>given</u> point in the space? What, if any, are the geometric features common to all such regions? The answers are found in the concepts of characteristic spheroids and characteristic ellipsoids. These objects will be studied in the third paper.

(d) At what time after the initial instant t_o will the n -ary scattered radiant energy be in steady state within the carrier space? What concept will determine that time t' after t_o at which the entire radiation field will be within a certain "distance" of its steady state configuration? These and related questions will be studied in the fourth paper dealing with temporal diameters.

(e) Finally, the question of the correct definition of time constants in bounded and unbounded spaces will be answered and illustrated in the fifth paper.

The present series, because of the inherent nature of the Divide and Conquer Principle, cannot exhaust all of the questions associated with the present approach. However, it is hoped that the five basic questions considered above will show the power of the approach and encourage further research in this direction.

Perhaps the greatest service that the present series can hope to perform is to emphasize and illustrate a fundamental thesis that all too few workers in radiative transfer and mathematical physics in general have come to realize, much less put into practice: the day of the exact abstract mathematical approach in theoretical physics is here, and here to stay. The economy of thought that this approach provides along with its far-reaching generalizations allow it to outstrip all but the most advanced of classical methods and endow it with power which cannot be dismissed or ignored with impunity.

EXTREMAL TIMES AND DISTANCES

Axiomatic Approach

In an earlier work[2] we introduced the so-called Transfer Process Axiom into the general mathematical framework of radiative transfer theory. This axiom embodies in abstract form the geometrical optics features of radiative transfer theory. In particular, the axiom supplies a collection $\{T_{t_1,t_2}\}$ of one to one measure preserving transformations T_{t_1,t_2} on the general carrier space $(X, \underset{\sim}{\Xi}, v)$ onto itself. If p_1 is a point of X, then the set

$$P(t_1,t_2) = \{ T_{t_1,t'}(p_1) : t_1 \leq t' \leq t_2 \}$$

is called the <u>natural path in</u> X determined by the point p_1 and time interval (t_1,t_2) (see, Radiative Measure Axiom, reference 1), where

$$p_i = T_{t_1,t_i}(p_1) \quad , \quad i = 1, 2 .$$

Conversely, suppose p_1 and p_2 are points in X on some natural path in X. Then there exist times t_0, t_1, and t_2 and a point p_0 such that

$$p_1 = T_{t_0,t_1}(p_0) \quad ,$$

$$p_2 = T_{t_0,t_2}(p_0) \quad .$$

This observation forms the basis for the following definition: The <u>temporal semimetric</u> t is a non negative function on $X \times X$ into the set of real numbers such that for $(p_1, p_2) \in X \times X$:

$$t(p_1, p_2) = |t_1 - t_2| \ ,$$

where t_1 , t_2 are associated with p_1 and p_2 on a natural path, in the manner described above.

In this way any two points p_1 , p_2 on a natural path in X are assigned a temporal distance between them. The <u>general</u> carrier space has no simple geometric structure. Hence no simple euclidean separation -- or distance -- in the usual sense can meaningfully be assigned to such pairs of points. However, by virtue of the Transfer Process Axiom, we can always assign a "temporal distance" between such points in a carrier space, and this distance is supplied by the temporal semimetric defined above.

By means of the axiomatized properties of the collection $\{T_{t_1, t_2}\}$ of elements of the Transfer Process and the introduction of the notion of temporal homogeneity $(T_{t_1, t_2} = T_{0, t_2 - t_1})$, t above possesses the usual properties of a semimetric as customarily defined in general metric space theory. In particular, whenever p_1 and p_2 are on a natural path, then temporal homogeneity implies

(i) $t(p_1, p_2) = t(p_2, p_1)$ (symmetry property)

and in general,

(ii) $t(p_1, p_2) = 0$ if and only if $p_1 = p_2$. (identity property)

Any non negative valued function t which has properties (i) and (ii), defines a <u>temporal semimetric</u> on X and X becomes a <u>temporal semimetric space</u>. If in addition, t has the property

$$\text{(iii)} \quad t(p_1, p_2) + t(p_2, p_3) \geq t(p_1, p_3), \quad \text{(triangle inequality)}$$

where p_1, p_2, and p_3 are <u>any</u> three points of X such that the pairs (p_1, p_2), (p_2, p_3), (p_1, p_3) are on natural paths, then t is called a <u>temporal metric</u> on X, and X becomes a <u>temporal metric space</u>.

General carrier spaces therefore become temporal semimetric spaces by introducing the temporal semimetric induced by the Transfer Process $\{T_{t_1, t_2}\}$. However, only very special kinds of carrier spaces are privileged to possess the status of a temporal metric space, i.e., a space in which t (as defined by the transfer process) satisfies (iii) in addition to the usual properties (i) and (ii). We shall conclude this paper with an example which illustrates this point. However, for the present we turn to an illustration of how a temporal semimetric may be obtained in the classical carrier spaces.

Fermat's Principle and Natural Paths

Let p' and p'' be two points in $(\Phi, \underset{\sim}{\Phi}, v)$, the classical carrier space (see Introduction of reference 1). Imbed p' and p'' in some continuous curve in Φ. If by some method we can constrain a light ray to travel along this curve between p' and p'', we would find that the time of travel $t(p', p'')$ between p' and p'' is given by:

$$t(p', p'') = \frac{1}{c} \int_P n(p)\, dr ,$$

where $n(p)$ is the index of refraction at p in X along the path P, and c is the velocity of light in a vacuum. The integral is over the path P, and r denotes geometrical length along P (as determined by the usual metric in Φ). If the factor $1/c$ is removed, the integral has the dimensions of length:

$$s(p', p'') = \int_P n(p)\, dr .$$

$s(p', p'')$ is called the optical length of P; this term is drawn from classical geometrical optics. Under suitable conditions there is only one natural path between p' and p''. The natural paths can be distinguished from all other paths between p' and p'' by the following criterion, which is one form of Fermat's Principle:

Let $\mathcal{P} = \{P_a : a \in A\}$ be the family of all continuous curves

P_a in Φ with ρ' and ρ'' as initial and terminal points. The index set A is some copy of E_n, $n \geq 1$. Then for each $a \in A$, there is a real number $T(a)$ defined by

$$T(a) = \int_{P_a} n\, dr .$$

Assuming T is differentiable on A, we then define the natural path $P(\rho', \rho'')$ between ρ' and ρ'' to be that $P_\nu \in \mathcal{P}$ with the property

$$\left. \frac{dT(a)}{da} \right|_{a=\nu} = 0 . \tag{1}$$

The Principle makes no assertion about the nature of the length of P_a, i.e., whether it is a local maximum or minimum. The statement of the Principle simply asserts that the function takes on an extreme value at $a = \nu$. To see whether $T(\nu)$ is a minimum or maximum we must examine the sign of

$$\left. \frac{d^2 T(a)}{da^2} \right|_{a=\nu} .$$

A particularly useful special case of the operation d/da described above is the well known "first variation" operation found in the calculus of variations. The first variation operation is excellently suited for engineering applications of the Principle in that it allows the Principle to be transformed into a set of partial differential equations for the natural path. The solutions of these equations are called extremals. The tracks of these solutions in Φ are the natural paths.

Classical Temporal Semimetrics

The natural path $P(p', p'')$ as found by the Fermat Principle has associated with it a real number $t(p', p'')$ defined by:

$$t(p', p'') = \frac{1}{c}\int_{P(p', p'')} n(p)\, dr ,$$

the <u>extremal time</u>. It follows immediately that the real valued function t on $\Phi \times \Phi$ as defined above has the properties (i) and (ii) of the temporal semimetric. t is called the <u>classical temporal semimetric</u>. In this way Φ becomes a temporal semimetric space. As is the case of the general carrier space, Φ is not necessarily a temporal metric space. The number

$$s(p', p'') = \int_{P(p', p'')} n(p)\, dr$$

is the <u>extremal distance</u> associated with $P(p', p'')$, which does not necessarily exist in general carrier spaces.

Euler Equations for the Extremals

Let F be a cartesian frame of reference for the location space component X of Φ . (Recall that Φ is the cartesian product of a location space (some well-defined subset of E_3) and the unit sphere in E_3). A curve in X may be defined by two continuous real valued functions y and z of x ; any point on the curve may thus be represented by the triple $(x, y(x), z(x))$ of real numbers.

Consider the integral

$$I(\rho',\rho'') = \int_{P(\rho',\rho'')} F(x,y,z,y',z')\,dx$$

of a function F along $P(\rho',\rho'')$. Here $y' = dy/dx$, and $z' = dz/dx$. The first variation of $I(\rho',\rho'')$ on the family $\mathcal{P}$ of all continuous curves between ρ' and ρ'' yields the following two (Euler) differential equations:

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left[\frac{\partial F}{\partial y'}\right] = 0\,,$$

$$\frac{\partial F}{\partial z} - \frac{d}{dz}\left[\frac{\partial F}{\partial z'}\right] = 0\,.$$

For the problem of determining the natural paths, we require F to be of the form:

$$F(x,y,z,y',z') = n(x,y,z)\left[1 + y'^2 + z'^2\right]^{1/2}.$$

With this F , the Euler differential equations become:

$$(1+y'^2+z'^2)\frac{\partial n}{\partial y} - n y'' = 0\,,$$

$$(1+y'^2+z'^2)\frac{\partial n}{\partial z} - n z'' = 0\,.$$

EXAMPLES

We now go on to consider some examples of the solution of the Euler equations, and conclude with an example which shows that the triangle inequality property of a metric is not in general satisfied by the extremal time function which defines the classical temporal semimetric.

Example 1

Let $(\Phi, \underset{\sim}{\Phi}, \nu)$ be a classical carrier space in which $n(\rho) = n$, a constant for all $\rho \in \Phi_0$, some subset of Φ. Then the Euler equations become:

$$y'' = 0 ,$$
$$z'' = 0 .$$

so that the pair of extremals

$$y = ax + b ,$$
$$z = cx + d .$$

define the general natural paths in Φ_0. It is clear that the general natural path in this case is a straight line (i.e., the track of a pair of <u>linear</u> functions of a single variable x).
It is easily shown that in this case

$$\left. \frac{d^2 T(a)}{da^2} \right|_{a=v} < 0 ,$$

so that the extremal times are minima.

Example 2

The second example may conceivably be of use in estimating the local geometric structure of natural paths in a <u>general</u> classical carrier space. Let $n(x,y,z) = n(z)$. In an arbitrary space this situation can be approximated locally by taking the gradient of the scalar point function n and then aligning the z-axis of the reference frame along the direction of the gradient.

The Euler equations now become:

$$y'' = 0 \quad ,$$

$$(1 + y'^2 + z'^2)\frac{dn}{dz} = n z'' \, ,$$

whence

$$y = ax + b \, ,$$

which shows that the natural paths are constrained to lie in the plane defined by $y = ax + b$.

The first integral of the other Euler equation is evaluated by setting $z' = \rho$. Then

$$(1 + a^2 + \rho^2)\frac{dn}{dz} = n\rho\frac{d\rho}{dz} \, ,$$

which is separable:

$$\frac{dn}{n} = \frac{\rho \, d\rho}{(1 + a^2 + \rho^2)} \quad ,$$

where a is the constant of integration of the first of the Euler equations. Integrating, we have

$$n = K(1 + a^2 + \rho^2)^{1/2}$$

where K is a constant of integration. From the Euler equation we have

$$\frac{(1+a^2+z'^2)}{z''} = \frac{n}{dn/dz}$$

so that

$$\frac{(1+a^2+z'^2)^{3/2}}{z''} = \frac{(1+a^2+z'^2)^{1/2}\, n}{dn/dz} .$$

The left side is precisely the radius of curvature $R(z, m)$ of the natural path through the general point (x, y, z), where $m = dz/dx$ evaluated at z. Therefore

$$R(z_0, m_0) = \frac{n^2(z_0)}{K(z_0)\,[dn/dz]} ,$$

where $K(z_0)$ is the value of the constant K for the natural path through (x_0, y_0, z_0), thus

$$K(z_0) = \frac{n(z_0)}{(1+a^2+m_0^2)^{1/2}}$$

There is nothing essential about the orientation of the plane $y = ax + b$; therefore, we are at liberty to choose $a = 0$. The formula for the radius of curvature becomes

$$R(z_0, m_0) = \frac{n(z_0)\,(1+m_0^2)^{1/2}}{[dn/dz]_{z_0}}$$

If any space satisfies the hypothesis of this example, the natural paths would then be plane curves. The torsion of such a curve is zero at each of its points. In general the analytical expression for the natural path can be determined given its torsion and radius of curvature at each point. Thus we can, if desired, reverse the present procedure and start with torsion and curvature conditions for $P(\rho', \rho'')$ and deduce its geometric features from its natural equations.

Example 3

This example shows why the triangle inequality cannot be expected to hold even in the most commonplace of classical carrier spaces. Figure 1 represents a region of a classical carrier space which has an index of refraction $n(p) = n \gg 1$ at all points except for the subregion X_0 defined by the (open set) rectangular parallelpiped outlined by but not including the points 1234, in which $n(\rho) = n > 1$ at each point ρ. The dimensions of X_0 are 2 x 2 x 2a units of length (in the usual metric of the space). It is easy to verify that a natural path AA_1A_2C of length 4a exists between A and C and is a straight line. The other natural paths AB, BC which are also straight lines (Example 1) are chosen such that $d(A,B) = d(B,C)$ and such that AB meets 2, and BC meets 3. Now using d, the triangle inequality (which holds for d) yields in particular:

$$d(A,B) + d(B,C) > d(A,C).$$

It will be shown that the corresponding inequality for the temporal distances does not hold. That is, we will show in particular that

$$t(A,B) + t(B,C) < t(A,C) .$$

Now $t(A,B) = d(A,B)/c$, and $t(B,C) = d(B,C)/c$, where c is the speed of light associated with an index of refraction $n = 1$. Further

$$t(A,C) = \frac{d(A,A_1)}{c} + \frac{d(A_1,A_2)}{(c/n)} + \frac{d(A_2,C)}{c} ,$$

therefore

$$t(A,C) = (n+1)\, 2a/c$$

Finally

$$t(A,B) + t(B,C) = 4(1+a^2)^{1/2}/c .$$

Hence

$$\frac{t(A,B) + t(B,C)}{t(A,C)} = \frac{2(1+a^2)^{1/2}}{a(n+1)} .$$

An inequality is obtained if, for example $a = 2$ and $n = 1.3$. This is sufficient to show that the triangle inequality for t does not hold in general; there are, nevertheless, an infinite number of pairs (a,n) for which the triangle inequality does hold.

The breakdown of the triangle inequality is essentially the consequence of a non-constant index of refraction function on X_o. It will be shown in a subsequent study that an arbitrary temporal semimetric space can be imbedded in a temporal metric space by the adoption of the concept of least local epoch time. In this way the powerful battery of theorems associated with metric space theory will be made accessible to the study of temporal metric spaces.

REFERENCES

1. Gottschalk, W. H., "Minimal Sets: An Introduction to Topological Dynamics," Bull. Am. Math. Soc. 64, 336-351 (1958).

2. Preisendorfer, R. W., "A Mathematical Foundation for Radiative Transfer Theory," J. Math. and Mech. 6, 685-730 (1957).

3. Preisendorfer, R. W., "Time-Dependent Principles of Invariance," Proc. Nat. Acad. Sci. 44, 328-332 (1958).

4. Ambarzumian, V. A., "On the Problem of the Diffuse Reflection of Light," J. Phys. Acad. Sci. U.S.S.R., 8, 65-75 (1944).

5. Chandrasekhar, S., Radiative Transfer, Oxford Press (1950).

6. Case, deHoffman, Placzek, Introduction to Neutron Diffusion Theory, Los Alamos Scientific Laboratory (1953) Vol. 1.

7. Riesz, F., and Sz. - Nagy. B., Functional Analysis, Ungar (1955).

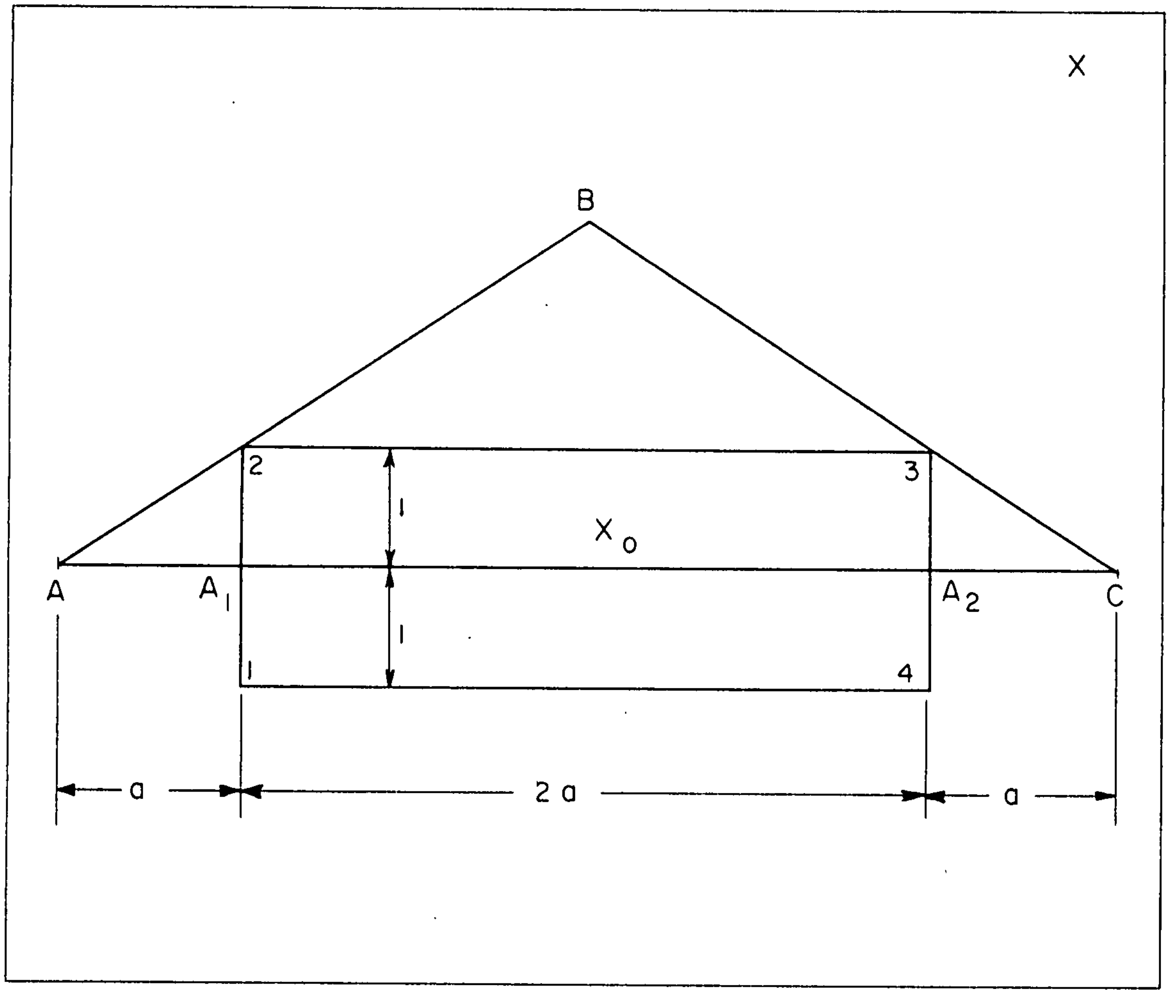

Figure 1

Rudolph W. Preisendorfer