Title
Markov chains and radiative transfer

Author
Preisendorfer, Rudolph W

Publication Date
1959-12-01

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

MARKOV CHAINS AND RADIATIVE TRANSFER

Rudolph W. Preisendorfer

December 1959
Index Number NS 714-100

Bureau of Ships
Contract NObs-72092

SIO Reference 59-72

Approved:

Seibert Q. Duntley

Seibert Q. Duntley, Director
Visibility Laboratory

Approved for Distribution

Roger Revelle

Roger Revelle, Director
Scripps Institution of Oceanography

Markov Chains and Radiative Transfer

Rudolph W. Preisendorfer

Scripps Institution of Oceanography, University of California

La Jolla, California

INTRODUCTION

This paper concludes the current series on radiative transfer theory in discrete spaces [1-7] by showing how a connection can be established between the theories of radiative transfer and Markov chains.

The series so far has been concerned with the internal affairs of the discrete theory: On the theoretical side, the details of formulation of the principle of local interaction were carried to the point where the fundamental equations for the light field were solved on an arbitrary discrete space X_n (reference 1). The discrete-space formulations were then used to derive the invariant imbedding relation which in turn yielded the principles of invariance on X_n (reference 2). In reference 3 the discrete-space formulations were extended to include the phenomenon of polarization. On the practical side of the internal affairs we have the explicit solution, ready for numerical evaluation, of the general two-flow equations on a linear lattice (reference 4). The practical utility of the discrete formulations was further underscored in reference 5 wherein the numerical computation procedure of reference 4 was extended to cover the twenty-six-flow problem in a cubic lattice representation of a plane-parallel optical medium (the type of medium which occurs in the study of light fields in planetary atmospheres and hydrospheres). Finally, in references 6 and 7, it was shown how the problems of the plane and point-source generated light fields could be formulated and solved on an arbitrary discrete space using the techniques developed in reference 1. These solutions can, in particular, be directed toward the problems of bioluminescence in the sea, and auroral and night-glow phenomena in the atmosphere.

The present paper complements the preceding seven by concentrating on certain apparently external aspects of discrete-space theory: It is shown how the notions of discrete-space radiative transfer can explicitly and in great detail be linked to a wide class of physical concepts which at first may be expected to lie outside its usual domain of ideas. This link is forged by demonstrating that the key principle of discrete-space transfer theory, namely the local interaction principle, can be cast into a form which satisfies the axioms of a Markov chain. In this way the set of ideas within the discrete space theory are explicitly given membership in the large class of physical processes --- ranging from probabilistic processes in atomic theory, through the dynamics of biological mutations, up to the mechanics of stellar motions in galaxies --- all of which fall within the domain of Markov chains or stochastic processes in general.

Besides serving to place discrete-space radiative transfer theory in its proper perspective within the modern domain of physical concepts, the theory is now open to the many analytical tools available in the discipline of Markov chains. Further, in order to effect the connection with the main stream of physics, the local interaction principle underwent a generalization which now includes a description of the phenomenon of scattering with change in wavelength.

MARKOV CHAINS

If one places a Mexican jumping bean at random on a checkerboard fitted with an encircling wall to keep the bean on the board, we have the realization of a Markov chain: Let $E_1, \ldots, E_{64}$ be the squares of the board numbered from left to right, top to bottom of the board. At time $t = 0$, which we shall call the <u>initial</u> <u>time</u>, let $p_1^0, \ldots, p_{64}^0$ be the probability that the bean is initially at rest on square E_1, $\ldots, E_{64}$, respectively. Now suppose that at $t = 0$ the bean is actually in E_1. We agree to let time increase on the board by one unit for every jump executed by the bean. Suppose further that at $t = 1$, i.e., after its first jump, the bean lands in E_6. There is a probability for this event. Denote it by p_{16}. This number is found experimentally by repeatedly placing the bean on E_1 and observing the number n_{16} of times out of a total number n of trials that it lands in E_6. In general, we define p_{ij} as the probability that the bean lands in E_j on the condition that it started its hop from E_i. We assume p_{ij} is independent of the time of the experiment. It is quite clear that we require of the numbers p_i^0 that

$$\sum_{i=1}^{64} p_i^0 = 1 ,$$

and that for each fixed i, $1 \le i \le 64$, we require that

$$\sum_{j=1}^{64} p_{ij} = 1$$

These, then are the essential ingredients of a Markov chain: A space $\mathcal{E}$ on which certain determinate or indeterminate events take place; a set of numbers p_i^o which measure the initial probability of an event on each point E_i of $\mathcal{E}$, and a set of numbers p_{ij} which give a measure of how the events at a point E_i may influence or give rise to an event at a generally different point E_j of $\mathcal{E}$. Formally, we have the:

Definition*: A Markov chain is an ordered triple $(\mathcal{E}, P^o, \mathcal{P})$ in which:

(i) $\mathcal{E} = (E_1, \dots, E_n)$ is a finite**, ordered set of elements of some set $\mathcal{E}$.

(ii) P^o is a non-negative valued function on $\mathcal{E}$ such that its values $P^o(E_i) \equiv p_i^o$ have the property $\sum_{i=1}^{n} p_i^o = 1$.

(iii) $\mathcal{P}$ is a non-negative valued function on $\mathcal{E} \times \mathcal{E}$ such that its values $\mathcal{P}(E_i, E_j) \equiv p_{ij}$ have the property: $\sum_{j=1}^{n} p_{ij} = 1$ for each i, $1 \le i \le n$.

New Markov Chains From Old

The conditions (ii) and (iii) of the definition endow a Markov chain with a useful regenerative property in the following sense.

---

* For a simple exposition of the notion of a Markov chain, see reference 7.

** We shall not need to consider infinite sets in this work.

Let $p_i^1 = \sum_{j=1}^{n} p_j^0 p_{ji}$, and $p_{ij}^2 = \sum_{k=1}^{n} p_{ik} p_{ki}$. Then the functions P^1 and $\mathcal{P}^2$ these relations define, have the property that $\sum_{i=1}^{n} p_i^1 = 1$ and $\sum_{i=1}^{n} p_{ij}^2 = 1$, $1 \leq j \leq n$, so that the ordered triple $(\mathcal{E}, P^1, \mathcal{P}^2)$ is a Markov chain.

In general, by induction, the ordered triple $(\mathcal{E}, P^n, \mathcal{P}^{n+1})$ is a Markov chain, where, symbolically

$$P^n = P^{n-1} \cdot \mathcal{P}$$

and

$$\mathcal{P}^{n+1} = \mathcal{P}^n \cdot \mathcal{P} \,,$$

or specifically,

$$p_i^n = \sum_{j=1}^{n} p_j^{n-1} p_{ji} \,, \qquad 1 \leq i \leq n \tag{1}$$

$$p_{ij}^{n+1} = \sum_{k=1}^{n} p_{ik}^n p_{ki} \,, \qquad 1 \leq i, j \leq n \,. \tag{2}$$

With these preliminaries in mind, we now may consider the details of establishing the connections between the local interaction principle and Markov chains.

PRELIMINARY CONNECTION

In an earlier work[9] a study was made of the mathematical foundations of radiative transfer which automatically subsumed both the continuous and discrete-space contexts. The radiative, and radiative transfer processes postulated in that study were related on a very general level to the notions of stochastic processes. It follows that the present results are in principle subsumed by the earlier analysis. However, for the explicit case of discrete carrier spaces such as those considered in the current series of investigations, it is an instructive and non trivial task to draw these connections in some detail.

We can give the reader a quick, preliminary insight into the nature of the connection between Markov chains and radiative transfer theory by using the point of view adopted in reference 1. There it was observed that radiative transfer theory is essentially the study of the ordered triple (X, N, T), where X is a space, N is the radiance function defined on X, and T is an operator (equation of transfer or local interaction principle) defined on N. We can now make the following pairing of concepts to establish the first, tentative, connection:

MARKOV CHAIN	RADIATIVE TRANSFER PROCESS
$(\mathcal{E}, P^{\circ}, \mathcal{P})$	(X, N, T)

$$\mathcal{E} \longleftrightarrow X$$

$$P^{\circ} \longleftrightarrow N$$

$$\mathcal{P} \longleftrightarrow T$$

Thus if we replace the Mexican jumping bean by a photon, and the checkerboard $\mathcal{E}$ with an optical medium X (the phase space) then P° the initial probability function pairs off with the initial radiance function N, and the transition probability function $\mathcal{P}$ pairs off with the transfer equation T.

The exact details of this pairing will now be examined for the case of an arbitrary discrete space to which the local interaction principle is applicable.

LOCAL CONNECTIONS

We begin the discussion by considering the initial local scattering function $\Sigma^{o}(x_i ; \cdot ; \cdot)$, and the initial local radiance vector $N^{o}(x_i)$ associated with an arbitrary fixed point x_i in a given discrete space X_n to which the local interaction principle is applicable. Let $\Xi_i = \{\xi_{i1}, \ldots, \xi_{in}\}$ be the local direction space at x_i and let $\Xi^{o}_i = \{\xi^{o}_{1i}, \ldots, \xi^{o}_{\Delta_i i}\}$ be the local source direction space at x_i. Let $\mathcal{E}_i = \Xi^{o}_i \cup \Xi_i$. Then $\mathcal{E}_i$ will be the event space for the Markov chain we are constructing. The function $\Sigma^{o}(x_i ; \cdot ; \cdot)$ is the candidate for the transition probability function $\mathcal{P}^{o}$ and the source radiance vector $N^{o}(x_i) = \{N^{o}_{1i}, \ldots, N^{o}_{\Delta_i i}\}$ is the candidate for the initial probability function P^{o}. Thus we consider the set of ordered triples $(\mathcal{E}_i , N^{o}(x_i), \Sigma^{o}(x_i ; \cdot ; \cdot))$, $1 \le i \le n$ for possible membership in the family of all Markov chains.

According to condition (ii) of the Markov definition, we require of the vector $N^{o}(x_i)$ the property that $\sum_{j=1}^{\Delta_i} N^{o}_{ji} = 1$, which of course can be endowed by fiat. However, the remaining condition (iii) on the transition function requires that :

$$\sum_{\xi \in \Xi_i} \Sigma(x_i ; \xi^{o} ; \xi) = 1$$

for each $\xi^{o} \in \Xi^{o}_i \subset \mathcal{E}_i$. This condition is not generally satisfied in non-conservative spaces. Or, putting it positively: (iii) is satisfied in conservative spaces only. This shortcoming on the part of $\overset{\circ}{\Sigma}(x_i ; \cdot ; \cdot)$ can be remedied by steering the following course: Recall first that the local conservation property[1] states:

$$\sum_{\xi \in \Xi_i} \Sigma^{\circ}(x_i ; \xi^{\circ} ; \xi) + A(x_i ; \xi^{\circ}) , \quad \xi^{\circ} \epsilon\ \Xi_i^{\circ} .$$

By defining $\Sigma^{\circ}(x_i ; \xi^{\circ} ; \xi_{\omega}) \equiv A(x ; \xi^{\circ}) , \quad \xi^{\circ} \epsilon\ \Xi_i^{\circ}$ (4)

and formally adjoining to Ξ_i the element ξ_{ω}, and to

$N^{\circ}(x_i)$ the corresponding formal component $N^{\circ}_{\omega i} = 0$, we have:

$$\sum_{\xi \epsilon \Xi_i'} \Sigma^{\circ}(x_i ; \xi^{\circ} ; \xi) = 1 \tag{5}$$

where Ξ_i' is the augmented local direction space, Ξ_i and where we have set $\Sigma^{\circ}(x_i ; \xi_{\omega} ; \xi_{\omega}) = 0$ by convention. Hence $\Sigma^{\circ}(x_i ; \cdot ; \cdot)$ on $\mathcal{E}_i \times \mathcal{E}_i$ satisfies condition (iii) and we have established that the ordered triple $(\mathcal{E}_i , N^{\circ}(x_i) , \Sigma^{\circ}(x_i ; \cdot ; \cdot))$ is a Markov chain for each i, $1 \leq i \leq n$.

GLOBAL CONNECTION

The global connection is established by welding together, in matrix form, the local connections established in the preceding section. The only change needed --- besides the assembling of larger vectors and matrices, is the renormalization of the components of N^o. That is, $N^o = [N^o(x_1), \dots, N^o(x_n)]$ is required to have its $\sum_{i=1}^{n} \Delta_i$ components add up to unity (see reference 1). Thus if we set $\mathcal{E} = (\mathcal{E}_1, \dots, \mathcal{E}_n)$ then the triple $(\mathcal{E}, N^o, \Sigma^o)$ is a Markov chain where N^o and Σ^o are given in their global forms established in reference 1.

We may verify that the ordered triple $(\mathcal{E}, N_+^o, M\Sigma)$ is a Markov chain where $N_+^o = N^o \Sigma^o$, M is a permutation matrix, and Σ is the scattering matrix for X_n augmented just as Σ^o was to include the local absorption function in its rows and columns. The domains of definition of its elements are Cartesian products of the form $(-\Xi_i') \times (\Xi_i')$, where the prime denotes the augmented space. The regenerative properties of a Markov chain (i.e., properties (1) and (2) above) allow the conclusion that $(\mathcal{E}, N_+^n, (M\Sigma)^{n+1})$ is also a Markov chain, where $N_+^n = N_+^o (M\Sigma)^n$.

CONNECTION VIA TRANSPECTRAL SCATTERING

At one stage (equation (4)) in establishing the local connections, a seemingly artificial step was taken to realize condition (ii) of the Markov chain definition. The artificiality is removed by suitably extending the concept of the local scattering function. As it stands now, $\Sigma(x_i; \cdot ; \cdot)$ is defined on $(-\Xi_i) \times (\Xi_i)$ or on $\Xi_i^0 \times \Xi_i$ in the case of $\Sigma^0(x_i, \cdot ; \cdot)$. Thus Σ describes the radiometric activity only in the local direction space at x_i. The function $A(x_i;\cdot)$ takes into account the remainder of the radiometric activity at x_i, namely scattering without change in wavelength to directions outside of Ξ_i; scattering with change in wavelength at x_i, and general absorption brought about by the conversion of radiant to non radiant energy.

All these diverse radiometric activities at each point of X_n can be described by a single function of the Σ-type. The domain of definition of the <u>local transpectral scattering function</u> at x_i is the product space $(-\Psi_i)\times(\Psi_i)$, where $\Psi_i = \{(\xi,\lambda) : \xi \epsilon \Xi_i , \lambda \epsilon \Lambda_i\}$, $-\Psi_i = \{(-\xi,\lambda) : \xi_i \epsilon \Xi_i , \lambda \epsilon \Lambda_i\}$, and $\Lambda_i = \{\lambda_{i1}, \ldots, \lambda_{i\ell_i}\}$ is the <u>local spectrum</u> at x_i. The set Λ_i represents the discrete version of the usual continuous electromagnetic wavelength domain $\Lambda = \{\lambda : 0 \leq \lambda \leq \infty\}$.

In the interests of generality we will allow the possibility of a different local spectrum set at each point of X_n. For most practical applications, however, we merely need $\Lambda_i = \Lambda_1$ for all i. Each number λ_{i_j} in Λ_i is associated with an actual wavelength in Λ, except for $\lambda_{i\ell_i}$. The latter denotes a radiant energy sink: The "place" where the radiant energy goes when it no longer is radiant energy as such (kinetic energy of thermal motion, potential energy levels of atoms and molecules, etc.).

The local transpectral scattering function $\Sigma(x_i; \cdot; \cdot)$ defined on $(-\Psi_i)\times(\Psi_i)$ at $x_i \varepsilon X_n$, with value $\Sigma(x_i; \psi'; \psi)$, denotes the fractional amount of the incoming (or field) radiance $N(x_i, \xi')$ that is "scattered" into the state ψ at x_i, where $\psi' = (\xi', \lambda')$ and $\psi = (\xi, \lambda)$.

The transpectral scattering function contains the monochromatic local scattering functions, implicitly defined for a fixed wavelength $\lambda \varepsilon \Lambda_i$, as a special case:

$$\Sigma(x_i; \xi'; \xi) = \Sigma(x; (\xi', \lambda); (\xi, \lambda)) \tag{6}$$

Furthermore, the local absorption function, implicitly defined for a fixed wavelength $\lambda \varepsilon \Lambda_i$, is now representable as:

$$A(x_i, \xi) = \sum{}' \Sigma(x_i; (\xi, \lambda); (\xi', \lambda')), \tag{7}$$

where $\sum{}'$ denotes summation over all $\xi' \varepsilon \Xi_i$, and over all $\lambda' \varepsilon \Lambda_i - \lambda$. Thus, the local conservation property becomes:

$$\sum_{\psi' \varepsilon \Psi_i} \Sigma(x_i; \psi; \psi') = 1, \quad \psi \varepsilon -\Psi_i \tag{8}$$

Finally, the transpectral form of the local interaction principle

becomes:

$$N(x_i, \psi_i) = \sum_{\psi_k \in (-\Psi_i)} \sum_{x_k \in X_n}' N(x_k, \psi_k)\, \Sigma(x_i ; \psi_k ; \psi_j')$$

$$+ \sum_{\psi_k^\circ \in (-\Psi_i^\circ)}' N^\circ(x_i, \psi_k^\circ)\;\; \Sigma^\circ(x_i ; \psi_k^\circ ; \psi_j), \tag{9}$$

where Ψ_i° is defined analagously to Ψ_i, except now we use Ξ_i°, instead of $-\Xi_i$. If we set $\sum_{i,k}' N^\circ(x_i, \psi_k^\circ) = 1$, $\mathcal{E}_i = (-\Psi_i) \cup (\Psi_i)$ then the triple $(\mathcal{E}_i, N^\circ, \Sigma^\circ)$ is once again a Markov chain, which may lead to higher order and global forms of Markov chains in precisely the manner shown in the preceding section.

CONCLUSION AND PROSPECTUS

We have shown how the study of radiative transfer processes on discrete spaces may legitimately be carried on as a subdiscipline of the theory of Markov chains by showing how the principle of local interaction can be cast into Markov chain form. As a consequence, the rich storehouse of results on Markov chains in general is now available for use in the study of old and new theoretical questions in radiative transfer. The possibility of combining the idea of nets of discrete spaces (reference 2) with the Markov chain formulation of radiative transfer theory opens a whole new field of exploration. Further, the study of asymptotic radiance distributions, time-dependent problems, etc., can now be formulated and viewed in novel perspectives by means of ergodic theory and recurrence phenomena; old random walk interpretations of radiative transfer phenomena take on greater depth; and the concept of radiative transfer processes finds its rightful place among the modern concepts of general stochastic processes. So once again the conceptual barriers between the events of radiative transfer theory and those of the main stream of physics are further dissolved, and we begin to observe what Mach long ago had discerned, namely that:

"Every event belongs, in a strict sense, to all the departments of physics, the latter being separated only by an artificial classification, which is partly conventional, partly physiological, and partly historical."

<u>The Science of Mechanics</u>, Open Court <u>Pub. Co.</u>, (La Salle, Ill., 1942, 9th Ed.) p. 597.

REFERENCES:

1. Preisendorfer, R. W., Radiative Transfer on Discrete Spaces, SIO Ref. 59-53, University of California, La Jolla, California, August 1959.

2. ----------------- Principles of Invariance on Discrete Spaces, SIO Ref. 59-55, University of California, La Jolla, California, August 1959.

3. ----------------- Theory of Polarized Light Fields in Discrete Spaces, SIO Ref. 59-66, University of California, La Jolla, California, November 1959.

4. ----------------- Radiative Transfer on a Linear Lattice, SIO Ref. 59-67, University of California, La Jolla, California, November 1959.

5. ----------------- Radiative Transfer on a Cubic Lattice, SIO Ref. 59-68, University of California, La Jolla, California, December 1959.

6. ----------------- Plane-Source Generated Light Fields in Discrete Spaces, SIO Ref. 59-70, University of California, La Jolla, California, December 1959.

7. ----------------- Two Fundamental Methods of Solving Point-Source Problems in Discrete-Space Radiative Transfer Theory, SIO Ref. 59-71, University of California, La Jolla, California, December 1959.

8. Feller, W., Probability Theory and its Applications, John Wiley and Sons, New York (1950).

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