Journal of Quantitative Spectroscopy & Radiative Transfer 112 (2011) 566–576

Contents lists available at ScienceDirect

Journal of Quantitative Spectroscopy & Radiative Transfer

journal homepage: www.elsevier.com/locate/jqsrt

Stochastic radiative transfer in Markovian mixtures: Past, present, and future

Evgueni Kassianov [a,*], Dana Veron [b]

[a] Pacific Northwest National Laboratory, Richland, WA 99352, USA
[b] University of Delaware, Newark, DE 19716, USA

ARTICLE INFO

Keywords:
Markovian statistics
Stochastic radiative transfer
Broken clouds
Homogeneous and inhomogeneous models
Analytical averaging

ABSTRACT

The Markovian approach, originally suggested in the early 1900s, has widespread practical use in many of our present-day studies and allows one to build bridges between diverse research areas such as statistical physics, astronomy, and computational science. This overview takes a broad sweep of several important examples with the emphasis on the stochastic radiative transfer in a cloudy atmosphere. In particular, the overview (i) highlights important contributions made by Pomraning and Titov to the neutron and radiation transport theory in a stochastic medium with homogeneous statistics and (ii) illustrates that ideas and tools introduced by these two distinguished scientists are gaining increasing impact and recognition in the atmospheric science.

1. Introduction

The 20th century opened auspiciously with a significant achievement in probability theory. In a sequence of highly celebrated papers, starting in 1906, gifted Russian mathematician, A.A. Markov (1856–1922), generalized various limit laws established for independent random variables (e.g., the law of large numbers) to dependent ones [1]. The generalization allows a departure from independence by including the most recent available information. In other words, the best prediction of a future event depends on what happens "today," and any information from "past" events is irrelevant in such predictions. In the course of his work A.A. Markov so advanced the theory of dependent variables, now called Markov chains (processes), that it became applicable to many areas. "Conceptually, a Markov process is the probabilistic analogue of the processes of classical mechanics, where the future development is completely determined by the present state and is independent of the way in which the present state has developed" [2, p. 369]. Interestingly, the advance of such statistical processes was inspired by the internal needs of probability theory, and Markov never discussed potential applications of his approach to physical science. However, the appearance of Markov's works eventually stimulated numerous studies, where the Markovian approach has been successfully tailored to different research areas, and a wealth of new ideas outside probability theory was flowering remarkably due to the suggested framework of statistical dependence.

This paper outlines several important applications of the Markovian approach and highlights related accomplishments with the main focus on stochastic radiative transfer (RT). Sections 2 and 3 give the basic information about Markov processes and fields, respectively. A general formulation of Markovian cloud fields is introduced in Section 4. Section 5 discusses application of these fields in the stochastic RT, and Section 6 includes some examples of the corresponding RT calculations. The conclusion is drawn in Section 7.

* Corresponding author. Tel.: +1 509 372 6535; fax: +1 509 372 6168.
E-mail addresses: Evgueni.Kassianov@pnl.gov (E. Kassianov), dveron@udel.edu (D. Veron).

0022-4073/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jqsrt.2010.06.011

2. Markov chain

Let us consider a sequence of random variables $\{X_t\}_{t\geq 0}=\{X_0, X_1,\ldots\}$ taking values in a discrete set of states $S=\{x_0, x_1,\ldots\}$, and define $X_t=X(t)$. A sequence of states $\{x_0, x_1,\ldots\}$ at successive times can be considered as a realization of a Markov chain. The simplest model assumes that random variables $\{X_0, X_1,\ldots\}$ are independent and identically distributed. However, such an assumption does not seem realistic for a wide range of physical phenomena. A natural step away from this assumption is the introduction of statistical dependences between variables at nearby times.

A family of random processes that describes such dependencies is called *Markov chains*. Specifically, a Markov chain is a random vector $\{X_0, X_1,\ldots\}$ for which the conditional probability satisfies the *Markov property*

$$P(X_{j+1}=x_{j+1}|X_i=x_i\ \ \forall i\leq j)=P(X_{j+1}=x_{j+1}|X_j=x_j). \tag{1}$$

In other words, the Markov property says that the conditional distribution of a future state, given the entire history of the process (up to time t_j), depends only the present state. The right-hand side of Eq. (1) is referred to as the *transition* probability $p_{ij}(t_i, t_j)$ from x_i to x_j. If the transition probability depends on a time difference $p_{ij}(t_i, t_j)=p_{ij}(t_i-t_j)$, the Markov chain is said to be *time-homogeneous*. Below we consider homogeneous Markov chains unless specifically stated otherwise. If times t_j are defined as an arithmetic progression $t_j=j\Delta t$, then for each $\Delta t>0$ the sequence of random variables $X(j\Delta t)$ represents a discrete-time Markov chain with a one-step transition probability $p_{ij}=p_{ij}(\Delta t)$.

The Poisson process is an example of a continuous-time Markov chain where $X(t)$ defines the total number of events that have occurred in the interval $[0,\ t)$. This counting process has stationary and independent increments with *exponentially* distributed intervals between consecutive events

$$p_{jk}(t-s)=P(X(t)=k|X(s)=j)=\frac{(\lambda(t-s))^{k-j}}{(k-j)!}\exp(-\lambda(t-s)), \tag{2}$$

where $0\leq s<t,\ \ j\leq k$ are non-negative integers and $k-j=0,\ 1,\ 2,\ldots$. The Poisson process provides an important framework for modeling multi-dimensional Markov random fields (Section 3). Note that several natural processes, including radioactive disintegration, can be described quite accurately by Poisson statistics under certain conditions, such as when the period of observations is smaller than the half-life of the short-lived isotopes [3].

Here is an example of the way in which Markov demonstrated his method. A piece of text can be treated as a sequence of random variables with two states $S=\{vowel,\ consonant\}$. Assuming that a sequence of letters is a chain, the corresponding probabilities can be obtained. To illustrate that, Markov considered two large samples of text. The first sample includes the sequences of 20,000 letters from Pushkin's novel in verse "Eugene Onegin" (a classic of Russian literature). The second sample includes the sequences of 100,000 letters from Aksakov's novel "Childhood of Bagrov the Grandson" (a major text in the Russian pastoral tradition). Markov performed calculations (by hand!) and illustrated that while the probabilities of vowels occurring were comparable (0.432 and 0.449), the probabilities of a vowel following a vowel diverged considerably (0.128 and 0.552). These striking results demonstrated that a piece of text can be viewed as a collection of dependent entities and that the language/author can be specified quite accurately when the relationship between these entities is identified. The Markovian approach provides a powerful tool for describing mathematically the communication phenomenon, and this tool has been applied successfully in different language-related studies, such as cryptography [4].

Perhaps the best-known early application of Markov chains goes back to the paper by Metropolis et al. [5], which was aimed at effectively computing multi-dimensional integrals with Boltzmann weights by using a random walk (Markov chain). Note that Metropolis "researched nuclear reactors with Enrico Fermi and Edward Teller, and, through his work with such noteworthy scientists, he came to the attention of J. Robert Oppenheimer, head of the Manhattan Project—the United States government's plan to build the first atomic bomb" [6, p. 254]. Hastings [7] introduced a more general version of the Metropolis algorithm and illustrated its performance with other popular one- and multi-dimensional distributions. The Metropolis–Hastings algorithm generates samples from a given distribution as the stationary solution of a suitably constructed Markov chain. A detailed description of the algorithm is given by Chib and Greenberg [8]. Since the Metropolis–Hastings algorithm is extremely versatile and allows one to simulate desired complex distributions, it has been used extensively in physics, computational science, and image analysis [9]. The methods based on simulating a Markov chain are applied widely for solving a system of simultaneous linear equations and integral equations as well [10].

3. Markov random fields

Markov random fields (MRF) are a natural extension of Markov chains on spatially distributed random variables. Markov chain formulation can be generalized to a random field setting in a straightforward manner, depending only on the definition of neighborhood structure. Specifically, a discrete MRF is a random vector $\{X_0, X_1,\ldots\}$ for which the conditional probability satisfies the local Markov property

$$P(X_i=x_i|X_j=x_j,\ j\neq i)=P(X_i=x_i|X_j=x_j,\ j\in\partial(i)), \tag{3}$$

where $\partial(i)$ is a set of neighbors of i. For example, the first order neighborhood includes two (left and right) or four (left, right, up and down) nearest sites with the same edge for one- and two-dimensional models, respectively.

The celebrated Hammersley–Clifford theorem states [11] that the MRF is a Gibbs Random field (GRF) in relation to a neighborhood $\partial(i)$ if the probability distribution function is given by

$$P(X)=\frac{1}{Z}\exp(-U_{\partial(i)}(X)), \tag{4}$$