# Probability and its Applications

*A Series of the Applied Probability Trust*

*Editors:* J. Gani, C.C. Heyde, T.G. Kurtz

**Springer**
*New York*
*Berlin*
*Heidelberg*
*Barcelona*
*Budapest*
*Hong Kong*
*London*
*Milan*
*Paris*
*Santa Clara*
*Singapore*
*Tokyo*

# Probability and its Applications

*Anderson:* Continuous-Time Markov Chains.
*Azencott/Dacunha-Castelle:* Series of Irregular Observations.
*Bass:* Diffusions and Elliptic Operators.
*Bass:* Probabilistic Techniques in Analysis.
*Choi:* ARMA Model Identification.
*Galambos/Simonelli:* Bonferroni-type Inequalities with Applications.
*Gani (Editor):* The Craft of Probabilistic Modelling.
*Grandell:* Aspects of Risk Theory.
*Gut:* Stopped Random Walks.
*Guyon:* Random Fields on a Network.
*Kallenberg:* Foundations of Modern Probability.
*Last/Brandt:* Marked Point Processes on the Real Line.
*Leadbetter/Lindgren/Rootzén:* Extremes and Related Properties of Random Sequences and Processes.
*Nualart:* The Malliavin Calculus and Related Topics.
*Rachev/Rüschendorf:* Mass Transportation Problems. Volume I: Theory.
*Rachev/Rüschendorf:* Mass Transportation Problems. Volume II: Applications.
*Resnick:* Extreme Values, Regular Variation and Point Processes.
*Shedler:* Regeneration and Networks of Queues.
*Todorovic:* An Introduction to Stochastic Processes and Their Applications.

Svetlozar T. Rachev Ludger Rüschendorf

# Mass Transportation Problems

## Volume II: Applications

Springer

Svetlozar T. Rachev
Department of Statistics
University of California
Santa Barbara, CA 93106
USA

Ludger Rüschendorf
Institut für Mathematische Stochastik
University of Freiburg
79014 Freiburg
Germany



Library of Congress Cataloging-in-Publication Data
Rachev, S.T. (Svetlozar Todorov)
Mass transportation problems / Svetlozar T. Rachev, Ludger Rüschendorf.
p. cm. — (Probability and its applications)
Includes bibliographical references and indexes.
Contents: vol. 1. Theory — vol. 2. Applications.
ISBN 0-387-98350-3 (vol. 1 : hardcover). — ISBN 0-387-98352-X (vol. 2 : hardcover)
1. Transportation problems (Programming) I. Rüschendorf, Ludger, 1948- . II. Title. III. Series: Springer series in statistics. Probability and its applications.
QA402.6.R33 1998
519.7′2—dc21 97-34593



ISBN 0-387-98352-X Springer-Verlag New York Berlin Heidelberg SPIN 10646816

*To my wife Zoja*
*and*
*To my parents Nadezda*
*and Todor Rachevi.*

Svetlozar (Zari) Rachev

*To my wife Gabi.*

Ludger Rüschendorf

# Preface to Volume II

The second volume of the *Mass Transportation Problems* is devoted to applications in a variety of fields of applied probability, queueing theory, mathematical economics, risk theory, tomography, and others. In Volume I we encompassed the general mathematical theory of mass transportation, concentrating our attention on:

- the general duality theory of the transportation and transshipment problem;
- explicit optimality results;
- applications to minimal probability metrics, stochastic ordering, approximation and extension problems;
- applications to functional analysis and mathematical economics (the Debreu theorem, utility theory, dynamical systems, choice theory, and convex and nonconvex analysis were dicsussed in this context).

In Volume II we expand the scope of applications of mass transportation problems. Some of them arise from modifications of the admissible transportation plans. In fact, for applications to mathematical economics it is of interest to consider relaxations of the marginal constraints, such as upper or lower bounds on the supply and demand distributions, or additional constraints like capacity bounds for the transportation plans. In mathematical tomography the basic problem is to reconstruct the multivariate probability distribution based on some information about the marginal distributions in a certain finite number of directions. This information may be represented by additional constraints on the support functions or distributional moments, or it may be contained in only partial information on the marginals. Thus there is a close relationship between a class of problems in mathematical tomography and the classical theory on moment problems, which again can be viewed as a relaxation on the set of constraints in mass transportation problems. We discuss in detail applications to approximation problems for stochastic processes and to rounding problems based on moment-type characteristics. A particular example will be the approximation of queueing models. The minimal metrics allow us to compare various rounding rules and to determine optimal ones from an asymptotic point of view.

An important field of applications of mass transportation problems we shall consider in this second volume is to probabilistic limit theorems. This approach was introduced in the seventies by the Russian school of probability theory, headed by V.M. Zolotarev. By inherent regularity properties of probability metrics defined via certain mass transportation problems, there are streamlined proofs for central limit theorems on Banach spaces yielding sharp quantitative estimates of Berry–Esseen type for the convergence rate. The probability metric approach will be applied to general stable and operator stable limits theorems, martingale-type limit theorems, limit behavior of summability methods, and compound Poisson approximation. A particular application is to the classical problem in mathematical risk theory dealing with sharp approximation of the individual risk model by the collective risk model. The probability metric approach will also be applied to the quantitative asymptotics in rounding problems. A new field of application of probability metrics arising as solutions of mass transportation problems is the analysis of deterministic and stochastic algorithms. This research area is of increasing importance in computer science and various fields of stochastic modeling. Based on regularity properties of probability metrics, a general "contraction" method for the asymptotic analysis of algorithms has been developed. The contraction method has been applied successfully to a variety of search, sorting, and other tree algorithms. Furthermore, the recursive structure in iterated functions systems (image encoding), fractal measures, bootstrap statistics, and time series (ARCH) models has been analyzed by this method. It becomes clear that there are many interesting probabilistic applications of this method to be rigorously developed in the future.

In the final chapter we consider applications to stochastic differential equations (SDEs) and to convergence of empirical measures. SDEs will be interpreted as continuous recursive structures. From this point of view we provide a detailed discussion on the approximative solution of nonlinear stochastic differential equations of McKean–Vlasov type by interactive particle systems with application to the Kac theory of chaos propagation. The probability metrics approach allows us to establish approximation results for various modifications of the diffusion system, some of them of "nontraditional" type. In a general context we establish approximation results for empirical measures and give applications to the approximation of stochastic processes. As final applications we discuss a weak approximation of SDEs of Itô type by a combination of the time discretization methods of Euler and Milshtein with a chance discretization based on the strong invariance (embedding) principle. This approximation is given in terms of minimal $L^p$-metrics and thereby based on regularity properties of the solutions of the corresponding mass transportation problem.

# Preface to Volume I

The subject of this book, mass transportation problems (MTPs), concerns the optimal transfer of masses from one location to another, where the optimality depends upon the context of the problem. Mass transportation problems appear in various forms and in various areas of mathematics and have been formulated at different levels of generality. Whereas the continuous case of the transportation problem may be cast in measure-theoretic terms, the discrete case deals with optimization over generalized transportation polyhedra. Accordingly, work on these problems has developed in several separate and independent directions.

The aim of this monograph is to investigate and to develop, in a systematic fashion, the *Monge–Kantorovich mass transportation problem (MKP)* and the *Kantorovich–Rubinstein transshipment problem (KRP)*. We consider several modifications of these problems known as the MTP with partial knowledge of the marginals and the MTP with additional constraints (MTPA). We also discuss extensively a variety of stochastic applications. In the first volume of *Mass Transportation Problems* we concentrate on the general mathematical theory of mass transportation. In Volume II we expand the scope of applications of mass transportation problems.

In 1781 Gaspard Monge proposed in simple prose a seemingly straightforward problem of optimization. It was destined to have wide ramifications. He began his paper on the theory of “clearings and fillings” as follows:

> When one must transport soil from one location to another, the custom is to give the name *clearing* to the volume of the soil that one

> must transport and the name *filling* ("remblai") to the space that it must occupy after transfer.
>
> Since the cost of transportation of one molecule is, all other things being equal, proportional to its weight and the interval that it must travel, and consequently the total cost of transportation being proportional to the sum of the products of the molecules each multiplied by the interval traversed; given the shape and position, the clearing and the filling, it is not the same for one molecule of the clearing to be moved to one or another spot of the filling. Rather, there is a certain distribution to be made of the molecules from the clearing to the filling, by which the sum of the products of molecules by intervals travelled will be the least possible, and the cost of the total transportation will be a *minimum*. (Monge, (1781, p. 666)).

In mathematical language Monge proposed the following nonlinear varational problem. Given two sets $A, B$ of equal volume, find an optimal volume-preserving map between them; the optimality is evaluated by a cost function $c(x, y)$ representing the cost per unit mass for transporting material from $x \in A$ to $y \in B$. The optimal map is the one that minimizes the total cost of transferring the mass from $A$ to $B$. Monge considered this problem with cost function equal to the Euclidean distance in $\mathbb{R}^d$: $c(x, y) = |x - y|$. Monge's problem turned out to be the prototype for a class of problems arising in various fields such as mathematical economics, functional analysis, probability and statistics, linear and stochastic programming, differential geometry, information theory, cybernetics, and matrix theory. The optimization function $\int_A c(x, t(x))\, \mathrm{d}x$ is nonlinear in the transportation function $t$, and moreover, the set of admissible transportations is a nonconvex set. This explains why it took a long time until even existence results for optimal solutions could be established. The first general existence result was given in 1979 by Sudakov.

On the second page of his paper Monge himself had remarked that to obtain a minimum, the intervals traversed by two different molecules should not intersect. This simple observation applied to the discrete case—where there are only a finite number of molecules—leads to a "greedy" algorithm, the so-called northwest corner rule. The totality of mass transferences plans in the discrete case is a polytope that arises in the transportation problem of mathematical programming, where it is treated in specialized form as an assignment problem and in generalized form as a network-flow problem. The northwest corner rule solves transportation problems having a particular structure on the costs and is, moreover, at the heart of many seemingly different problems having an "easy" solution (cf. Hoffman (1961), Barnes and Hoffman (1985), Derigs, Goecke, and Schrader (1986), Hoffman and Veinott (1990), Olkin and Rachev (1991), and Rachev and Rüschendorf (1994); see also Burkard, Klinz, and Rudolf (1994) and the references therein).

The Academy of Paris offered a prize for the solution of Monge's problem, which was claimed by the differential geometer P. Appell (1884–1928), who established some geometric properties of optimal maps in the plane and in $\mathbb{R}^3$. But it took a long time until a real breakthrough in the transportation problem came, originating in the seminal 1942 paper of L.V. Kantorovich entitled "On the transfer of masses." Kantorovich stated the problem in a new, abstract, and in more easily accessible setting and without knowledge of Monge's work. Kantorovich learned of Monge's work only later (cf. his 1948 paper). In the Kantorovich formulation of the mass transportation problem (the so-called "continuous" MTP), the initial mass (the clearing) and the final mass (the filling) can be considered as probability measures on a metric space. The essential step in this formulation is the replacement of the class of transportation map by the wider class of generalized transportation plans, that are identifiable with the convex set of all probability measures on the product space with fixed marginals. The difficult nonlinear Monge problem was thereby replaced by a linear optimization problem over an abstract convex set. This made it possible to put this problem in the framework of linear optimization theory and encouraged the development of general duality theory for the solution of the Kantorovich formulation of the transportation problem as the basic tool. Accordingly, these problems and their generalizations will be referred to as *Monge–Kantorovich Mass Transportation Problems (MKPs).*

Kantorovich's measure theoretic formulation made the problem accessible to various areas of the mathematical sciences and other scientific fields. Kantorovich himself received a Nobel Prize in Economics for related work in mathematical economics.(1) Here is a list of some references in the mathematical sciences:

- Functional analysis: Kantorovich and Akilov (1984)
- Probability theory: Fréchet (1951), Cambanis et al. (1976), Dudley (1976, 1989), Kellerer (1984), Rachev (1991c), Rüschendorf (1991)
- Statistics: Gini (1914, 1965), Hoeffding (1940, 1955), Kemperman (1987), Huber (1981), Bickel and Freedman (1981), Rüschendorf (1991)
- Linear and stochastic programming: Hoffman (1961), Barnes and Hoffman (1985), Anderson and Nash (1987), Burkard, Klinz and Rudolf (1994)
- Information theory and cybernetics: Wasserstein (1969), Gray et al. (1975), Gray and Ornstein (1979), Gray et al. (1980)
- Matrix theory: Lorentz (1953), Marcus (1960), Olkin and Pukelsheim (1982), Givens and Shortt (1984)

(1) L.V. Kantorovich together with T.C. Koopmans received the Nobel Memorial Prize in Economic Science in 1975 for "contributions to the theory of optimum allocation of resources"; see Dudley (1989, p. 342).

Many practical problems arising in various scientific fields have led mathematicians to solve MKPs: e.g., in

- Statistical physics: Tanaka (1978), Dobrushin (1979)
- Reliability theory: Barlow and Proschan (1975), Kalashnikov and Rachev (1990), Beneš (1985)
- Quality control: Jirina and Nedoma (1957)
- Transportation: Dantzig and Ferguson (1956)
- Econometrics: Shapley and Shubik (1972), Pyatt and Round (1985), Gretsky, Ostroy, and Zame (1992)
- Expert systems: Perez and Jirousek (1985)
- Project planning: Haneveld (1985)
- Optimal models for facility location: Ermoljev, Gaivoronski, and Nedeva (1983)
- Allocation policy: Rachev and Taksar (1992)
- Quality usage: Rachev, Dimitrov and Khalil (1992)
- Queueing theory: Rachev (1989), Anastassiou and Rachev (1992a, 1992b)

There are several surveys in the vast literature about MKP, among them Rachev (1984b), Rachev and Rüschendorf (1990), Burkard, Klinz, and Rudolf (1994), Cuesta-Albertos, Matrán, Rachev, and Rüschendorf (1996), and Gangbo and McCann (1996) related to dual solutions and applications of MKP; Shorack and Wellner (1985, Sect. 3.6) on optimal processes; Benes and Stepan (1987, 1991) on extremal mass transportation plans; Rüschendorf (1981, 1991, 1991a), Kellerer (1984), Rachev (1991c) on multivariate transportation problems; Dudley (1989) on distances in the space of measures; Talagrand (1992) and Yukich (1991) on matching problems.

In recent years, characterizations of the solutions of the Monge–Kantorovich problem have been given in terms of $c$-subgradients of generalized convex functions defined in terms of the cost functions $c(x, y)$ (cf. Knott and Smith (1984, 1992), Brenier (1987), Rüschendorf and Rachev (1990), Rüschendorf (1991, 1991a, 1995), Cuesta-Albertos, Matrán, Rachev, and Rüschendorf (1996), and Gangbo and McCann (1996)).

For the case of squared Euclidean costs $c(x, y) = |x-y|^2$, the generalized convexity property is equivalent to convexity, and $c$-subgradients are identical to the usual subgradients of convex analysis. From this characterization a series of explicit solutions of the transportation problem could be established. It also implies that the solutions of the MKP are under continuity assumptions given by mappings. Therefore, the solutions of the "easier" MKP imply as well the existence and characterizations of solutions of the original Monge problem, and so the MKP turns out to be the fundamental formulation of the transportation problem. For this reason, we concentrate in this book on the Kantorovich-type mass tranportation problems. For a discussion of interesting analytic aspects of the Monge problem, we refer to Gangbo and McCann (1996).

Another type of MTP appears in probability theory, even if it leaves the framework of probability measures as transportation plans. Its solutions are bounded measures on a product of two spaces with the difference of marginals equal to the difference of two given probability measures. It will be called the *Kantorovich–Rubinstein Problem (KRP)*, since the first results were obtained by Kantorovich and Rubinstein (1958). In its relation to the practical task of mass transportation it is sometimes referred to as the transshipment problem; see Kemperman (1983), and Rachev and Shortt (1990). The KRP has been developed to a great extent in the Russian school of probabilists and functional analysts, in particular by V.L. Levin, A.A. Milyutin, and A.M. Vershik and their students.

For metric cost functions the KRP coincides with the corresponding MKP; for general cost functions it can be reduced to the MKP for a corresponding reduced cost function. For the duality theory of the KRP a specific detailed theory with many results that are of value in themselves has been developed with wide-ranging applications to mathematical economics. For a different approach to the KRP as introduced in Dudley (1976) and as further extended in Rachev and Shortt (1990) we refer to the book of Rachev (1991c).

A problem related to both MKP and KRP is the *Mass Transportation Problem with Partial Knowledge of the Marginals (MTPP)*, which is expressed by stating finitely many moment conditions. Problems of this type were formulated and extensively studied by Rogosinski (1958), Kemperman (1983), and Kuznezova-Sholpo and Rachev (1989). Barnes and Hoffman (1985) considered mass tranportaion problems with capacity constraints on the admissible transportation plans as an example of *Mass Transportation Problems with Additional Constraints (MTPA)* (see Rachev (1991b) and Rachev and Rüschendorf (1994)).

In this book we give an extensive account of the duality theory of the MKP and the KRP, including the known results on explicit constructions and characterizations of optimal solutions.

In Chapters 2 and 3 we present important duality theorems for the Monge–Kantorovich problem based on work of H. Kellerer, L. Rüschendorf, S.T. Rachev, and D. Ramachandran.

In Chapters 4 and 5 we present basically work of V.L. Levin; we analyze measure-theoretic methods for infinite-dimensional linear programs developed in context with the KRP as well as applications to general utility theorems (the Debreu theorem), extension theorems, choice theory, and set-valued dynamical systems.[(2)]

In Chapters 6 and 8 we discuss new material on applications of the MKP and the KRP to the representation of ideal metrics and on various probabilistic approximation and limit theorems. This supplements the earlier results in this direction as described in the book of Rachev (1991) on probability metrics and stochastic models. In particular, we show that probability metrics allow us to find unified proofs for central limit theorems for martingales, (operator) stable limit theorems, and to more specific problems like compound Poisson approximation or rounding problems.

Chapter 7, the first chapter in the second volume, is concerned with modifications of the MKP by additional or relaxed constraints. We discuss various types of moment problems and applications to the tomography paradoxon and to the approximation of queueing systems. A wide range of applications of metrics based on the transportation problem has been established in recent years in connection with recursive stochastic equations. We discuss algorithms of informatics (sorting, searching, branching, search trees) as well as applications to the approximation of stochastic differential equations, to the propagation of the chaos property of particle systems with applications to the approximation of nonlinear PDEs, as well as to the rate of convergence of empirical measures, which is of interest for matching problems in Chapters 9 and 10.

From the technical point of view, MKPs can be subdivided into the discrete and continuous cases, according to the nature of their basic spaces and to the supports of the initial and the final masses. In the discrete case, the totality of the mass transference plans is the polytope that arises in the transportation problem of mathematical programming. There is, of course, a vast literature on the transportation problem, its specialization to the assignment problem, and its generalization to network flow problems. It turns out, as will be elaborated further in the book, that the northwest corner rule in the discrete case corresponds to a closed form for the solution in the continuous case. Indeed, the discrete analogue of a result known in the continuous case provides a new result in the discrete case; and its simple proof in the discrete case provides a new proof for the continuous case, see Rachev and Rüschendorf (1994c) and the references therein. Another approach in the discrete linear case prefers to exploit the special structure of supplies and demands (or clearings and fillings) and permits a particularly simple combinatorial algorithm for finding an optimal solution as developed

[(2)] These two chapters were written following closely the notes kindly provided to us by V.L. Levin.

by Balinski (1983), Balinski and Russakoff (1984), Balinski (1985, 1986), Goldfarb (1985), Kleinschmidt, Lee, and Schannath (1987), and Burkard, Klinz, and Rudolf (1994).

MTPs may be viewed as an analogue and a unifying framework of a problem considered by probabilists at the beginning of the twentieth century: *How does one measure the difference between two random quantities?* Many specific contributions to the analysis of this problem have been made, including Gini's (1914) notion of concordance, Kendall's $\tau$, Spearman's $\varrho$, the analysis of greatest possible differences by Hoeffding (1940) and others, by Fréchet (1951, 1957), Robbins (1975), and Lai and Robbins (1976), and the generalizations of these results by Cambanis, Simons, and Stout (1976), Rüschendorf (1980), Tchen (1980), and Cambanis and Simons (1982). These (and others) offer piecemeal answers to basic questions that arise from different stochastic models; they give no guidance as to the question of what concept should be used where: There is no general theory underlying the diverse approaches. We refer to Kruskal (1958), Gini (1965), and Rachev (1984b, 1991c).

In this book we investigate, develop, and exploit the connections between the discrete and continuous versions of the mass transportation problems (MTP) as well as study systematically the relationships between the methods and results from different versions of the MTP. The MTPs are the basis of many problems related to the question of stability of stochastic models, to the question of whether a proposed model yields a satisfactory approximation to the phenomenon under consideration, and to the problem of approximation of stochastic and deterministic algorithms. It is our belief that MTPs hold great promise in stochastic analysis as well as in mathematical analysis. The MTP is full of connections with geometry, (partial) differential equations, (generalized) convex analysis, moment problems, infinite-dimensional linear programming, measurable choice theory, and extension problems, and it has many open problems. It has a great potential for a series of applications in several scientific fields.

This book grew out of joint work and lectures delivered by the authors at the Steklov Mathematical Institute, Universität Münster, Universität Freiburg, the Ecole Polytechnique, SUNY at Stony Brook, and the University of California, Santa Barbara, over many years. Many colleagues provided helpful suggestions after reading parts of the manuscript. All chapters were rewritten several times, and preliminary versions were circulated among friends, who eliminated many inaccuracies and obscurities. We would like to thank H.G. Kellerer, V.L. Levin, M. Balinski, D. Ramachandran, G.A. Anastassiou, M. Maejima, M. Cramer, I. Olkin, M. Gelbrich, W. Römisch, V. Beneš, L. Uckelmann, and many other friends and colleagues who encouraged us to complete the work. We are indebted to Mrs. M. Hattenbach and Ms. A. Blessing for their superb typing; the appearance of this monograph owes much to them. We are grateful to the publisher and especially to J. Kimmel for support and patience. We are particularly thankful to J. Gani for his invaluable suggestions concerning improvements of this work, his help with the organization of the material, and his encouragement to continue the project.

Finally, we thank the Alexander von Humboldt Foundation for its generous financial support of S.T. Rachev in 1995 and 1996, which made this joint work possible. [3]

---

(3) The work of S.T. Rachev was also partially supported by NSF Grants. The joint work of the authors was supported by NATO-Grant CRG900798.

# Contents to Volume II

# Contents to Volume I