THE CONCENTRATION OF MEASURE PHENOMENON

Michel Ledoux

Institut de Mathématiques
Université Paul-Sabatier
31062, Toulouse (France)

TABLE OF CONTENTS

INTRODUCTION

The aim of this book is to present the basic aspects of the concentration of measure phenomenon. The concentration of measure phenomenon was put forward in the early seventies by V. Milman in the asymptotic geometry of Banach spaces. Of isoperimetric inspiration, it is of powerful interest in applications, in various areas such as geometry, functional analysis and infinite dimensional integration, discrete mathematics and complexity theory, and especially probability theory. This book is concerned with the basic techniques and examples of the concentration of measure phenomenon with no claim to be exhaustive. A particular emphasis has been put on geometric, functional and probabilistic tools to reach and describe measure concentration in a number of settings.

As mentioned by M. Gromov, the concentration of measure phenomenon is an elementary, yet non-trivial, observation. It is often a high dimensional effect, or a property of a large number of variables, for which functions with small local oscillations are almost constant. A first illustration of this property is suggested by the example of the standard n-sphere $\mathbb{S}^n$ in $\mathbb{R}^{n+1}$ when the dimension n is large. One striking aspect of uniform measure σ^n on $\mathbb{S}^n$ in high dimension is that it is almost concentrated around the equator. More generally, as a consequence of spherical isoperimetry, given any measurable set A with, say, $\sigma^n(A) \geq \frac{1}{2}$, almost all points (in the sense of the measure σ^n) on $\mathbb{S}^n$ are within (geodesic) distance $\frac{1}{\sqrt{n}}$ from A (which becomes infinitesimal as $n \to \infty$). Precisely, for every $r > 0$,

$$\sigma^n(A_r) \geq 1 - \mathrm{e}^{-(n-1)r^2/2}$$

where $A_r = \{x \in \mathbb{S}^n ; d(x, A) < r\}$ is the neighborhood of order $r > 0$

of A for the geodesic metric on $\mathbb{S}^n$.

This concentration property on the sphere may be described equivalently on functions, an idea going back to P. Lévy. Namely, if F is a continuous function on $\mathbb{S}^n$ with modulus of continuity $\omega_F(\eta) = \sup\{|F(x) - F(y)|; d(x,y) < \eta\}$, then

$$\sigma^n\big(\big\{|F - m_F| \geq \omega_F(\eta)\big\}\big) \leq 2\,\mathrm{e}^{-(n-1)\eta^2}$$

where m_F is a median of F for σ^n. Therefore, functions on high dimensional spheres with small local oscillations are strongly concentrated around a mean value, and are thus almost constant on almost all the space! This high dimensional concentration phenomenon was extensively used and emphasized by V. Milman in his investigation of asymptotic geometric analysis.

As yet another interpretation, the concept of observable diameter as considered by M. Gromov is a "visual" description of the concentration of measure phenomenon. We view the sphere with a naked eye which cannot distinguish a part of $\mathbb{S}^n$ of measure (luminosity) less than $\kappa > 0$ (small but fixed). A Lipschitz function F may be interpreted as an observable, that is an observation device giving us the visual image measure of σ^n by F. In this language, Lévy's inequality on Lipschitz functions expresses that the "observable diameter of $\mathbb{S}^n$" is of the order of $\frac{1}{\sqrt{n}}$ as n is large, in strong contrast with the diameter of $\mathbb{S}^n$ as a metric space.

In probability theory, the concentration of measure is a property of a large number of varibales, such as in laws of large numbers. A probabilistic description of the concentration phenomenon goes back to E. Borel who suggested the following geometric interpretation of the law of large numbers for sums of independent random variables uniformly distributed on the interval $[0,1]$. Let μ^n be uniform measure on the n-dimensional cube $[0,1]^n$. Let H be a hyperplane that is orthogonal to a principal diagonal of $[0,1]^n$ at the center of the cube. Then, if H_r is the neighborhood of order $r > 0$ of H, for every $\varepsilon > 0$, $\mu^n(H_{\varepsilon\sqrt{n}}) \to 1$ as $n \to \infty$. Actually, the relevant observable is $x \in [0,1]^n \mapsto \frac{1}{n}\sum_{i=1}^n x_i \in [0,1]$ that concentrates around the mean value $\frac{1}{2}$. The normal projection of the cube to the principal diagonal identified with $[0,\sqrt{n}]$ thus sends most of the measure of the cube to the subsegment

$$\Big[\frac{\sqrt{n}}{2} - \varepsilon\sqrt{n}, \frac{\sqrt{n}}{2} + \varepsilon\sqrt{n}\Big].$$

In fact, $\varepsilon\sqrt{n}$ may be replaced by any sequence $r_n \to \infty$ as follows from the central limit theorem.

A related description is the following. Let $X_1, X_2, \ldots$ be a sequence of independent random variables taking the values ± 1 with equal probability, and set, for every $n \geq 1$, $S_n = X_1 + \cdots + X_n$. We think of S_n as a function of the individual variables X_i and we state the classical law of large numbers by saying that S_n is essentially constant (equal to 0). Of course, by the central limit theorem, the fluctuations of S_n are of order $\sqrt{n}$ which is hardly zero. But as S_n can take values as large as n, this is the scale at which one should measure S_n in which case S_n/n is indeed essentially zero as expressed by the classical exponential bound

$$\mathbb{P}\Big(\Big\{\frac{|S_n|}{n} \geq r\Big\}\Big) \leq 2\,\mathrm{e}^{-nr^2/2}, \quad r \geq 0.$$

In this context, and according to M. Talagrand, one probabilistic aspect of measure concentration is that a random variable that depends (in a smooth way) on the influence of many independent variables (but not too much on any of them) is essentially constant.

Measure concentration is surprisingly shared by a number of cases that generalize the previous examples, both by replacing linear functionals (such as sums of independent random variables) by arbitrary Lipschitz functions of the samples, and by considering measures that are not of product form. It was indeed again the insight of V. Milman to emphasize the difference between the concentration phenomenon and standard probabilistic views on probability inequalities and law of large number theorems by the extension to Lipschitz (and even Hölder type) functions and more general measures. His enthusiasm and persuasion eventually convinced M. Talagrand of the importance of this simple, yet fundamental, concept.

It will be one of the purposes of this book to describe some of the basic examples and applications of the concentration of measure phenomenon. While the first applications were mainly developed in the context of asymptotic geometric analysis, they have now spread to a wide range of frameworks, covering areas in geometry, discrete and combinatorial mathematics, and in particular probability theory. Classical probabilistic inequalities on sums of independent random variables have been used indeed over the years in limit theorems and discrete algorithmic mathematics. They provide quantitative illustrations of measure concentration by so-called exponential inequalities (mostly of Gaussian type). Recent developments on the concentration of measure phenomenon describe far reaching extensions that provide dimension free concentration properties in product spaces which, due to the work of M. Talagrand during the last decade, will form a main part of these notes.

The book is divided into 8 chapters. The first one introduces the notions and elementary properties of concentration functions, deviation inequalities and their more geometric counterparts as observable diameters. We also briefly indicate a few useful tools to investigate concentration properties. The second chapter describes some of the basic and classical isoperimetric inequalities at the origin of the concentration of measure phenomenon. However, we do not concentrate on the usually somewhat delicate extremal statements, but rather develop some self-contained convexity and semigroup arguments to reach the concentration properties originally deduced from isoperimetry. Chapter 3 is a first view towards geometric and topological applications of measure concentration. In particular, we describe there Milman's proof of Dvoretsky's theorem on almost spherical sections of convex bodies. V. Milman in this proof most vigorously emphasized the usefulness of concentration ideas. Chapter 4 investigates measure concentration in product spaces, mostly based on the recent developments by M. Talagrand. After a brief view of the more classical martingale bounded difference method, we cover there the convex hull and finite point approximations, which are of powerful use in applications to both empirical processes and discrete mathematics. We also discuss the particular concentration property of the exponential distribution. The next two chapters emphasize functional inequalities stable under products thereby obtaining a new approach to the results of Chapter 4. Chapter 5 is devoted to the entropic and logarithmic Sobolev inequality approach. We present there the Herbst method to deduce concentration from a logarithmic Sobolev inequality and describe the various applications to product measures and related topics. Chapter 6 is yet another form of concentration relying on information and transportation cost inequalities with which one may reach several of the conclusions of the preceding chapters. Chapter 7 is devoted to the probabilistic applications of concentration in product spaces to sharp bounds on sums of independent random vectors or empirical processes: these applications lay at the heart of M. Talagrand's original investigation. The last chapter is a selection of (recent) applications of the concentration of measure phenomenon to various areas such as statistical mechanics, geometric probabilities, discrete and algorithmic mathematics, for which the concentration ideas, although perhaps at some mild level, appear to be useful tools of investigation.

While we describe in this work a number of concentration properties put forward in several contexts, from more geometric to functional and probabilistic settings, we usually produce the correct orders bu almost never discuss sharp constants.

This book is strongly inspired by early references on the subject. In particular, the lecture notes by V. Milman and G. Schechtman that describe the concentration of measure phenomenon and its applications to asymptotic theory of finite dimensional normed spaces were a basic source of inspiration during the preparation of this book. We also used the recent survey by G. Schechtman in the Handbook in the Geometry of Banach Spaces. (The latter handbook contains further contributions that illustrate the use of concentration in various functional analytic problems.) The memoir by M. Talagrand on isoperimetric and concentration inequalities in product spaces is at the basis of most of the material presented starting with Chapter 4, and the ideas developed there gave a strong impetus to recent developments in various areas of probability theory and its applications. Several of the neat arguments presented in these references have been reproduced here. The already famous $3\frac{1}{2}$ Chapter of the recent book by M. Gromov served as a useful source of geometric examples where further motivating aspects of convergence of metric measure spaces related to concentration are developed. While many geometric invariants are introduced and analyzed there, our point of view is perhaps a bit more quantitative and motivated by a number of recent probabilistic questions. Perspectives and developments related to the concentration of measure phenomenon in various areas of mathematics and its applications are discussed in M. Gromov's book as well as in the recent papers of M. Gromov and V. Milman in the special issues "Vision" of Geometric and Functional Analysis GAFA2000.

Each chapter is followed by some Notes and Remarks with in particular attempt to trace the origin of the main ideas. We apologize for inaccuracies and omissions.

The notations used throughout this book are the standard ones used in the literature. Although we keep some consistency, we did not try to unify all the notations and often used the classical notation in a given context even though it might have been used differently

in another.

I am grateful to Michel Talagrand for numerous discussions over the years on the topic of concentration and for explaining to me his work on concentration in product spaces. Parts of several joint works with Sergey Bobkov on concentration and related matters are reproduced here. I sincerely thank him for corrections and comments on the first draft of the manuscript. I also thank Vitali Milman and Gideon Schechtman for their interest and useful comments and suggestions, and Markus Neuhäuser, James Norris and Vladimir Pestov for helpful remarks and corrections. I sincerely thank the A.M.S. Mathematics Editor Edward Dunne for his help in the preparation of the manuscript.

Toulouse, May 2001 Michel Ledoux