



Introduction to stochastic variational problems

Victor L. Berdichevsky

► To cite this version:

| Victor L. Berdichevsky. Introduction to stochastic variational problems. 2011. hal-00601764

HAL Id: hal-00601764

<https://hal.science/hal-00601764>

Preprint submitted on 20 Jun 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Introduction to stochastic variational problems

Victor L. Berdichevsky

Mechanical Engineering, Wayne State University, Detroit MI 48202 USA

Abstract The lectures provide an introduction to the Chapters on stochastic variational problems from the author's book Variational Principles of Continuum Mechanics, Springer, 2009.

0.1 Reminder from complex analysis

The major goal of these lectures is to explain how to compute probabilities in some stochastic variational problems. The lectures can be used as an introduction to the Chapters of my book Variational Principles of Continuum Mechanics, Springer-Verlag, 2009, which are concerned with stochastic variational problems. Before proceeding to stochastic variational problems we have to learn how to compute probability in much simpler cases, like, for example, how to find probability distribution of a sum of independent random variables. Unfortunately, these issues are discussed in engineering probability courses at the time when the students do not have enough mathematical background, and a simple and beautiful nature of the classical results of probability theory, like the central limit theorem or Gauss distribution, remain unrevealed. Therefore, I will spend the first part of the course to cover these issues.

To do the calculations we will use complex analysis. I would assume with a great deal of certainty, that not everyone in this room was exposed to properly taught complex analysis. In high school you, perhaps, were taught that the complex unity i is a square root from -1 :

$$i = \sqrt{-1}.$$

What does that mean? Of course, you can solve the equation $x^2 + 1 = 0$ in terms of i and write $x = \pm i$. So what? Well, you can then write the solution of any polynomial equation in the form $\alpha + i\beta$, where α and β are real numbers, besides, the number of roots is equal to the power of the polynomial. This is nice indeed. This was the way in which the complex numbers were introduced in XVI century by Italian mathematicians

Cardano and Bombelli. This way penetrated in modern text books without much change. Unfortunately, one key word is missing in such a treatment, the word without which the real understanding of complex numbers is hardly possible. I begin with an explanation of what the complex numbers are (note that complex numbers are not numbers!) and why they are needed for the problems under consideration. Besides, I will review the basics of complex analysis. Then we spend two lectures for classical results of probability theory, and then go on to stochastic variational problems.

Complex numbers. We know two basic operations with vectors: We can multiply vector, \vec{a} , by a number, λ ; if a^i ($i = 1, \dots, n$) are the components of a vector \vec{a} in some basis \hat{e}_i ¹,

$$\vec{a} = a^i \hat{e}_i,$$

then vector $\lambda \vec{a}$ has the components λa^i :

$$\lambda \vec{a} = \lambda a^i \hat{e}_i.$$

We also can sum vectors; if a^i and b^i are the components of vectors \vec{a} and \vec{b} , then vector $\vec{a} + \vec{b}$ has the components $a^i + b^i$:

$$\vec{a} + \vec{b} = (a^i + b^i) \hat{e}_i.$$

The latter definition corresponds to the parallelogram rule (Fig. 1).

Now we wish more. We wish to operate with vectors as we do with numbers. We wish to introduce multiplication of vectors, \vec{a} and \vec{b} , in such a way that the product of vectors is a vector, and, as for numbers,

$$\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}, \quad (\vec{a} + \vec{b}) \cdot \vec{c} = \vec{a} \cdot \vec{c} + \vec{b} \cdot \vec{c}, \quad (\vec{a} \cdot \vec{b}) \cdot \vec{c} = \vec{a} \cdot (\vec{b} \cdot \vec{c}).$$

Besides, we would like to be able to divide vectors, i.e., for each vectors \vec{a} and \vec{b} , we should be able to compute their ratio, a vector \vec{c} ,

$$\vec{c} = \vec{a} / \vec{b}.$$

For given \vec{a} and \vec{b} , vector \vec{c} must be determined uniquely from the equation

$$\vec{b} \cdot \vec{c} = \vec{a}.$$

¹In all formulas summation over repeated indices is assumed, e.g.

$$a^i \hat{e}_i \equiv \sum_{i=1}^n a^i \hat{e}_i.$$

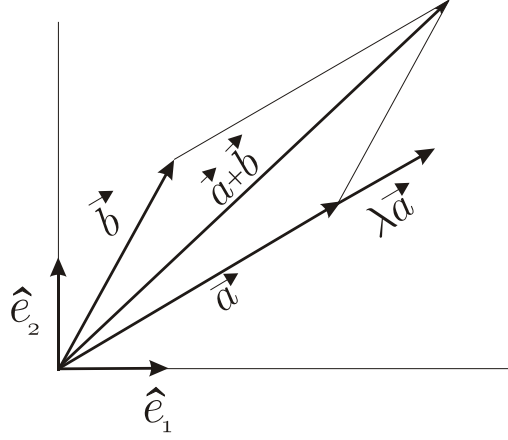


Figure 1. The definition of a sum of two vectors and multiplication of a vector by number.

Clearly, the products of vectors which are introduced in vector analysis, the vector product, $\vec{a} \times \vec{b}$, and the dot product, $\vec{a} \cdot \vec{b}$, do not fit: $\vec{a} \times \vec{b}$ is anti-symmetric ($\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$), the dot product is a scalar, not a vector. The multiplication operation we are looking for must be a new one. It turns out that such product can be introduced only in three spaces: two-dimensional, four-dimensional and eight-dimensional. The vectors of the corresponding two-dimensional space are called complex numbers, four-dimensional space quaternions and eight-dimensional space octonions. Quaternions were first conceived by W.R. Hamilton, octonions by A. Caley (Caley numbers). Only vectors of two-dimensional space, complex numbers, possess all the features of usual numbers; quaternions and octonions do not.

Now we have to define what is the product of two vectors in two-dimensional space. To this end, it is enough to define the products of the basic vectors. Then the product of any two vectors can be computed from the relation

$$(a^1 \hat{e}_1 + a^2 \hat{e}_2) \cdot (b^1 \hat{e}_1 + b^2 \hat{e}_2) = a^1 b^1 \hat{e}_1 \cdot \hat{e}_1 + (a^1 b^2 + a^2 b^1) \hat{e}_1 \cdot \hat{e}_2 + a^2 b^2 \hat{e}_2 \cdot \hat{e}_2.$$

Let us denote the two basic vectors by the symbols $\hat{e}_1 = \hat{1}$, $\hat{e}_2 = \hat{i}$. Then any vector z has the form $z = x \cdot \hat{1} + y \hat{i}$. The products of basic vectors are defined by the following rule:

$$\hat{1} \cdot \hat{1} = \hat{1}, \quad \hat{1} \cdot \hat{i} = \hat{i} \cdot \hat{1} = \hat{i}, \quad \hat{i} \cdot \hat{i} = -\hat{1}.$$

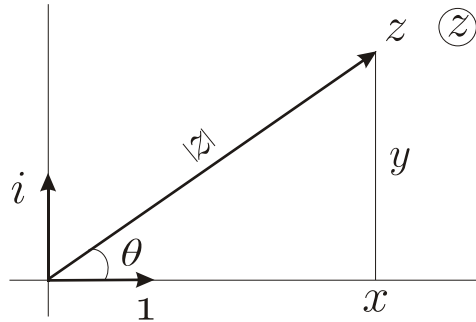


Figure 2. Notations in complex plane.

$$z_1 = x_1 \cdot \hat{1} + y_1 \hat{i}, \quad z_2 = x_2 \cdot \hat{1} + y_2 \hat{i},$$

their product is

$$z_1 \cdot z_2 = (x_1 x_2 - y_1 y_2) \hat{1} + (x_1 y_2 + x_2 y_1) \hat{i}.$$

One can prove that such a product has all the features of the product of usual numbers.

The tradition is not to write the hat at the basic vector \hat{i} : one writes for this basic vector $i \equiv \hat{i}$. Moreover, without confusion the basic vector $\hat{1}$ in the expression $x\hat{1} + iy$ can be dropped. So the complex number takes the form

$$z = x + iy.$$

Such form defines the components of vector z , x and y , uniquely. They are called real and imaginary parts of z and denoted by $Re z$ and $Im z$, respectively. The complex numbers with zero imaginary part are called real. The two-dimensional space with such defined vector product is called complex plane. We arrive at the usual definition of complex numbers, which you can find e.g. in Wikipedia: complex numbers are the numbers of the form $x + iy$, where x, y are real numbers while $i^2 = -1$. What is missing in this definition is the key word: vector. Complex numbers are not numbers, they are vectors.

E x e r c i s e s. Complex conjugate of $z = x + iy$ is, by definition, the number $\bar{z} = x - iy$. Show that $z\bar{z}$ is a real number equal to the squared length of vector z ; the length of z , $|z|$ is defined as $|z| = \sqrt{x^2 + y^2}$ (see Fig.2). Find the number $1/z$.

Functions of complex variables. An advantage we gain working in complex plane is that we can operate with vectors in the same way we do with numbers. In particular, all functions we used in calculus, make sense being applied to vectors. For example, for numbers, function e^x can be defined as a sum

$$e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \dots = \sum_{n=0}^{\infty} \frac{1}{n!}x^n.$$

In the same way it can be defined for vectors because the sum of vectors and the powers $z^n = \underbrace{z \dots z}_{n \text{ times}}$ are well defined:

$$e^z = 1 + z + \frac{1}{2}z^2 + \dots = \sum_{n=0}^{\infty} \frac{1}{n!}z^n. \quad (1)$$

Of course, one has to complement this by a notion of convergence, but this is done in a natural way: sequence $z_n \rightarrow 0$ as $n \rightarrow \infty$, if $|z_n| \rightarrow 0$ as $n \rightarrow \infty$. By multiplying two series,

$$\sum_0^{\infty} \frac{1}{n!}z_1^n \quad \text{and} \quad \sum_0^{\infty} \frac{1}{n!}z_2^n,$$

one can check that

$$e^{z_1+z_2} = e^{z_1}e^{z_2}. \quad (2)$$

An immediate consequence of (1) and (2) is Euler formula

$$e^{iy} = \cos y + i \sin y. \quad (3)$$

Indeed, plugging in (1) $z = iy$, and using that

$$i^n = \begin{cases} (-1)^k i & \text{if } n = 2k+1 \\ (-1)^k & \text{if } n = 2k, \end{cases}$$

we get

$$e^{iy} = \sum_{\text{even } n} \frac{1}{n!}z^n + \sum_{\text{odd } n} \frac{1}{n!}z^n = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}y^{2k} + i \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}y^{2k+1}. \quad (4)$$

Euler formula follows from the Taylor expansion of \sin and \cos :

$$\cos y = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}y^{2k}, \quad \sin y = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}y^{2k+1}.$$

Denote by θ the angle between the vector z and positive real axis. We use the convention, that angle θ increases in counter clockwise direction and changes within the limits $-\pi, \pi$. According to Euler formula one can write z in the polar form:

$$z = |z| e^{i\theta}. \quad (5)$$

Another important function of complex variable to be encountered further is $\ln z$. It is defined as inversion of exponential function, i.e.

$$e^{\ln z} = z. \quad (6)$$

Since, according to (5),

$$z = e^{\ln|z| + i\theta},$$

we can rewrite (6) as

$$e^{\ln z} = e^{\ln|z| + i\theta}$$

from which

$$\ln z = \ln|z| + i\theta. \quad (7)$$

Construction of functions by means of a series makes a class of functions defined by a converging Taylor series especially important; they are called analytic functions. More precisely, $f(z)$ is an analytic function in a region D if it can be presented in D by a converging series. We drop further all mathematical details, they can be found in numerous sources, and focus only on the basic ideas.

Derivative of a function of complex variable is defined in terms of its Taylor series²

$$f(z) = \sum_{k=0}^{\infty} a_k z^k,$$

as

$$f'(z) = \sum_{k=0}^{\infty} k a_k z^{k-1}.$$

This definition is equivalent to the usual one

$$f'(z) = \lim_{\Delta z \rightarrow 0} \frac{f(z + \Delta z) - f(z)}{\Delta z}$$

²Here we consider a function which is analytical in a vicinity of the point $z = 0$. If a function is analytical in a vicinity of some point z_0 , Taylor series is written with respect to the shift $z - z_0$:

$$f(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k.$$

provided the convergence of all series, which are involved.

Integral over a contour γ in z -plane,

$$\int_{\gamma} f(z) dz,$$

is defined in the same way as for a contour in two-dimensional real plane.

Integral of an analytic function over a closed contour in a simply-connected region³ is zero. Indeed, let a function be analytic in a vicinity of the point $z = 0$. The integral over closed contour of z^n is zero:

$$\oint_{\gamma} z^n dz = \oint_{\gamma} d \frac{z^{n+1}}{n+1} = 0.$$

Thus, integrating the Taylor series of an analytic function we get zero for any analytic function. Consideration of functions analytical in vicinity of non-zero points is similar.

An example of non-analytic function in vicinity of $z = 0$ is the function $f(z) = 1/z$; it is analytic in any region with excluded point $z = 0$ and non-analytic in any vicinity of point $z = 0$.

E x e r c i s e s. 1. Write down polar forms for $1/z$, \bar{z} , $1/\bar{z}$, -1 , z^n .

2. Find $(1+i)^5$, $(1+i)^{100}$.

3. Derive from the definitions of trigonometric functions of complex variables,

$$\sin z = \frac{1}{2i} (e^{iz} - e^{-iz}), \quad \cos z = \frac{1}{2} (e^{iz} + e^{-iz}),$$

that

$$\sin^2 z + \cos^2 z = 1.$$

4. Hyperbolic sin and cos are defined by the formulas:

$$\sinh z = \frac{1}{2} (e^z - e^{-z}), \quad \cosh z = \frac{1}{2} (e^z + e^{-z}).$$

Show that

$$\cosh^2 z - \sinh^2 z = 1.$$

5. Show that

$$\sin 2z = 2 \sin z \cos z.$$

³A region is simply-connected if any closed contour can be shrunk by a continuous transformation to a point.

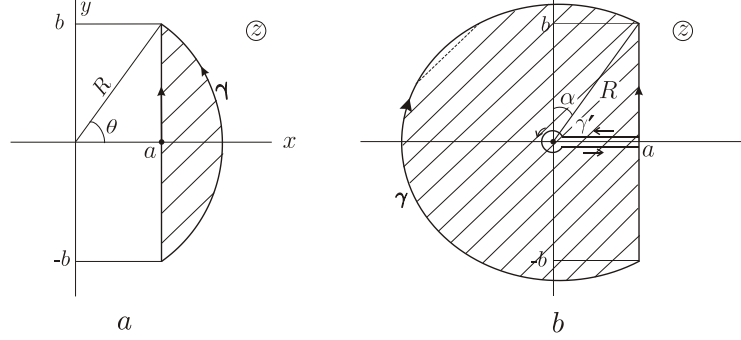


Figure 3. To justification of formula (10).

6. Show that for functions $f(z)$ with real coefficients of Taylor expansion

$$\overline{f(z)} = f(\bar{z}).$$

7. Write Taylor series for the function $(e^z - 1)/z$.

Step function. The only reason why we need complex numbers in what follows is the presentation of the step function by means of an integral in complex plane. By the step function, $\theta(t)$, we mean the following function

$$\theta(t) = \begin{cases} 1 & t > 0 \\ 0 & t < 0. \end{cases} \quad (8)$$

Usually, the value of this function at $t = 0$ is not essential. We will set for definiteness

$$\theta(0) = \frac{1}{2}. \quad (9)$$

We need the following integral presentation of $\theta(t)$:

$$\theta(t) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{e^{tz}}{z} dz, \quad a \geq 0, \quad (10)$$

where the integration contour is a straight line in the right half-plane z .

Since formula (10) will play a central role in further calculations, we outline here its derivation.

First of all we need to specify the meaning of the integral in (10). The integral has infinite limits, and, in general, one can write

$$\int_{a-i\infty}^{a+i\infty} \frac{e^{tz}}{z} dz = \lim_{b_1, b_2 \rightarrow \infty} \int_{a-ib_1}^{a+ib_2} \frac{e^{tz}}{z} dz. \quad (11)$$

Such a formula would be meaningful if the integral converges absolutely, i.e. there exist an integral

$$\int_{a-i\infty}^{a+i\infty} \left| \frac{e^{tz}}{z} \right| |dz|.$$

The role of absolute convergence follows from the inequality: for any contour γ ,

$$\left| \int_{\gamma} f(z) dz \right| \leq \int_{\gamma} |f(z)| |dz|.$$

This inequality is a consequence of the inequality

$$|z_1 + \dots + z_n| \leq |z_1| + \dots + |z_n|,$$

which obviously follows from the triangle inequality,

$$|z_1 + z_2| \leq |z_1| + |z_2|.$$

If an integral with infinite limits converges absolutely, it does not matter in which way the infinite limits are approached. The integral in (11) does not converge absolutely. Indeed

$$\int_{a-i\infty}^{a+i\infty} \left| \frac{e^{tz}}{z} \right| |dz| = \int_{-\infty}^{+\infty} \frac{|e^{t(a+iy)}|}{|(a+iy)|} dy = \int_{-\infty}^{+\infty} \frac{e^{ta}}{\sqrt{a^2 + y^2}} dy.$$

This integral diverges logarithmically. Therefore, we have to specify what is meant by integral in (11). We set

$$\int_{a-i\infty}^{a+i\infty} \frac{e^{tz}}{z} dz = \lim_{b \rightarrow \infty} \int_{a-ib}^{a+ib} \frac{e^{tz}}{z} dz. \quad (12)$$

Let us compute this limit. First of all, for $t = 0$

$$\theta(0) = \frac{1}{2\pi i} \lim_{b \rightarrow \infty} \int_{a-ib}^{a+ib} \frac{1}{a+iy} i dy = \frac{1}{2\pi i} \lim_{b \rightarrow \infty} (\ln(a+ib) - \ln(a-ib)).$$

From (7)

$$\ln(a+ib) = \ln \sqrt{a^2 + b^2} + i \arctan \frac{b}{a}$$

$$\ln(a - ib) = \ln \sqrt{a^2 + b^2} - i \arctan \frac{b}{a}.$$

Since

$$\lim_{b \rightarrow \infty} \arctan \frac{b}{a} = \frac{\pi}{2},$$

we have

$$\theta(0) = \frac{1}{2\pi} \left[\frac{\pi}{2} - \left(-\frac{\pi}{2} \right) \right] = \frac{1}{2}.$$

This is in accord with (9).

Let now $t < 0$. We aim to show that the limit in (12) is zero. Function e^{tz}/z is analytic in the shadowed region in Fig. 3a. Therefore, integral of e^{tz}/z over the boundary of the shadowed region is zero. Hence, the integral of e^{tz}/z over the straight segment $[a - ib, a + ib]$ is equal to the integral over the contour γ . Let γ be an arc of a circle with radius $R = \sqrt{a^2 + b^2}$ and the center at $z = 0$. Then the integral is bounded from above:

$$\left| \int_{\gamma} \frac{e^{tz}}{z} dz \right| \leq \int \frac{|e^{tz}|}{|z|} |dz| = \int_{-\theta}^{\theta} e^{tR \cos \theta} d\theta.$$

Since $t < 0$, this integral does not exceed the integral,

$$J(R) = \int_{-\pi/2}^{\pi/2} e^{-|t|R \cos \theta} d\theta = 2 \int_0^{\pi/2} e^{-|t|R \cos \theta} d\theta.$$

We are going to show that function $J(R)$ tends to zero as $R \rightarrow +\infty$. This is the first point in these lectures where we encounter the necessity to find the asymptotics of an integral; we will have several such problems later on. If the function in the exponent, $\cos \theta$, were strictly positive everywhere on the integration interval, i.e. $\cos \theta \geq \min \cos \theta = c > 0$, then the integrand does not exceed $e^{-|t|R \min(\cos \theta)}$, and the integral tends to zero as $R \rightarrow +\infty$. In our case, however, $\min \cos \theta = 0$, (see Fig. 4a), it is achieved at $\theta = \pi/2$.

We have to study the contribution to the integral of the vicinity of the point $\theta = \pi/2$. Let us split the integral in two parts

$$\frac{1}{2} J(R) = \int_0^{\pi/2 - \alpha} e^{-|t|R \cos \theta} d\theta + \int_{\frac{\pi}{2} - \alpha}^{\pi/2} e^{-|t|R \cos \theta} d\theta. \quad (13)$$

The first integral in (13) does not exceed

$$e^{-|t|R \cos(\frac{\pi}{2} - \alpha)} \left(\frac{\pi}{2} - \alpha \right)$$

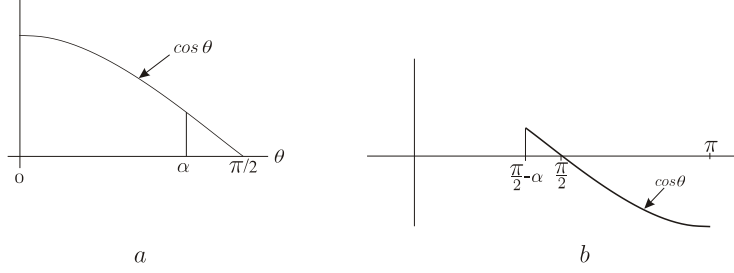


Figure 4. Graph of $\cos \theta$.

and tends to zero as $R \rightarrow +\infty$. For sufficiently small α , in the second integral we can replace $\cos \theta$ by the first terms of Taylor expansion

$$\cos \theta = \cos \theta \Big|_{\theta=\pi/2} + \frac{d \cos \theta}{d\theta} \Big|_{\theta=\pi/2} \left(\theta - \frac{\pi}{2} \right) = \left(\frac{\pi}{2} - \theta \right).$$

Then the second integral takes the form

$$\int_{\frac{\pi}{2}-\alpha}^{\pi/2} e^{-|t|R(\frac{\pi}{2}-\theta)} d\theta = \int_0^{\alpha} e^{-|t|R\xi} d\xi.$$

We can increase this integral replacing the upper limit by $+\infty$. Then we get the integral which is computed analytically

$$\int_0^{\infty} e^{-|t|R\xi} d\xi = \frac{1}{|t|R}.$$

Hence, $J(R) \rightarrow 0$ as $R \rightarrow +\infty$, and $\theta(t) = 0$ for $t < 0$ indeed.

Consider now the case $t > 0$. In this case we introduce a region shadowed in Fig. 3b. Function e^{tz}/z is analytic in a shadowed region, and integral over the boundary of the shadowed region is zero. Therefore, the integral over the segment $[a - ib; a + ib]$ is equal to the sum of integrals over contours γ and γ' , γ being a circle of radius R . The integral over γ can be estimated from above as

$$\left| \int_{\gamma} \frac{e^{tz}}{z} dz \right| = \int_{\gamma} \frac{|e^{tz}|}{|z|} |dz| = \int_{\frac{\pi}{2}-\alpha}^{\frac{3\pi}{2}+\alpha} e^{tR \cos \theta} d\theta = 2 \int_{\frac{\pi}{2}-\alpha}^{\pi} e^{tR \cos \theta} d\theta. \quad (14)$$

Here α is the angle shown in Fig. 3b. The integral in the right hand side of (14) can be written as a sum

$$\int_{\frac{\pi}{2}-\alpha}^{\pi} e^{tR \cos \theta} d\theta = \int_{\frac{\pi}{2}}^{\pi} e^{tR \cos \theta} d\theta + \int_{\frac{\pi}{2}-\alpha}^{\frac{\pi}{2}} e^{tR \cos \theta} d\theta.$$

Function $\cos \theta$ on the segment $[\pi/2 - \alpha, \pi]$ is shown in Fig. 4b. The first member of the sum coincides with $J(R)$ considered above and thus tends to zero as $R \rightarrow \infty$. The behavior of the second integral is not immediately clear. Let us estimate this integral using that $R \cos \theta \leq a$:

$$\int_{\frac{\pi}{2}-\alpha}^{\pi/2} e^{tR \cos \theta} d\theta \leq e^{ta} \int_{\frac{\pi}{2}-\alpha}^{\pi/2} d\theta = \alpha e^{ta}.$$

As follows from Fig. 3b, $\alpha \rightarrow 0$ as $R \rightarrow \infty$. Hence, the second integral tends to zero as well. So, as $b \rightarrow \infty$, the integral of e^{tz}/z over the segment $[a - ib; a + ib]$ tends to the integral over γ' , which, in turn, is equal to the integral over a circle of small radius r ,

$$\begin{aligned} \frac{1}{2\pi i} \oint \frac{e^{tz}}{z} dz &= \frac{1}{2\pi i} \int \frac{e^{tre^{i\theta}}}{re^{i\theta}} rie^{i\theta} d\theta \\ &= \frac{1}{2\pi} \oint e^{tre^{i\theta}} d\theta \Big|_{r \rightarrow 0} \rightarrow \frac{1}{2\pi} \oint d\theta = 1, \end{aligned}$$

as claimed.

Sometimes, it is convenient to have in formula (10) the integral over a line in the left half-plane. The corresponding relation is obtained from (10), if we notice that the integral of e^{tz}/z over the boundary of the shadowed region in Fig. 5 is zero.

Therefore, the integral of e^{tz}/z over the segment $[a - ib; a + ib]$ is equal to the sum of integrals over $[-a - ib; -a + ib]$, γ^+ , γ^- and γ' . The integral over γ^+ tends to zero as $b \rightarrow \infty$:

$$\left| \int_{\gamma^+} \frac{e^{tz}}{z} dz \right| \leq \int_{-a}^a \frac{|e^{tz}|}{|z|} dx = \int_{-a}^a \frac{e^{tx}}{\sqrt{x^2 + b^2}} dx \leq e^{ta} \frac{1}{b} \cdot 2a.$$

Similarly, integral over γ^- tends to zero. Integral over γ' is equal to 1. Hence,

$$\theta(t) = 1 + \frac{1}{2\pi i} \int_{-a-i\infty}^{-a+i\infty} \frac{e^{tz}}{z} dz, \quad a \geq 0. \quad (15)$$

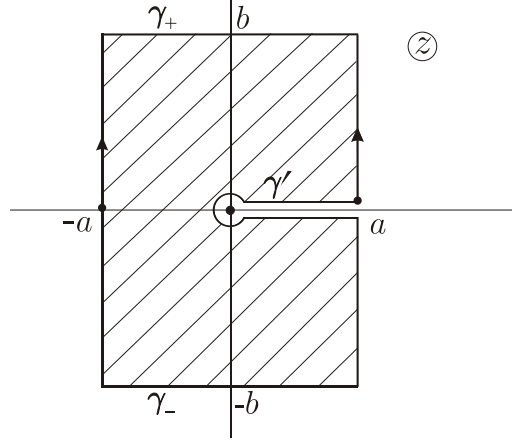


Figure 5. To the move of integration contour to the left half-plane.

It is known, that derivative of the step function is δ -function:

$$\frac{d\theta(t)}{dt} = \delta(t). \quad (16)$$

Differentiating (10) over t we formally obtain a presentation of δ -function

$$\delta(t) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} e^{tz} dz. \quad (17)$$

This relation, strictly speaking, does not make sense, because the integral in (17) does not converge. However, δ -function is used only in the form of integrals with smooth functions. In such cases, one can write

$$\int \delta(t) \varphi(t) dt = \frac{1}{2\pi i} \int_{a-ib}^{a+ib} \int \varphi(t) e^{tz} dt dz, \quad (18)$$

and, if the function of z ,

$$\int \varphi(t) e^{tz} dt,$$

decays fast enough as $|z| \rightarrow \infty$, (18) holds true.

E x e r c i s e s. 1. Let $f(z)$ be an analytic function in a region D . Show that for a point $z \in D$ and for any contour C , surrounding z ,

$$f(z) = \frac{1}{2\pi i} \int_C \frac{f(\zeta) d\zeta}{\zeta - z}.$$

2. Let $f(z) = u(x, y) + iv(x, y)$ be an analytic function. Show that real and imaginary part of f , u and v , are linked by a system of equations (Cauchy-Riemann equations)

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}$$

$$\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

3. Derive from Cauchy-Riemann equations that $u(x, y)$ and $v(x, y)$ are harmonic functions, i.e.

$$\Delta u(x, y) = 0, \quad \Delta v(x, y) = 0,$$

where Δ is Laplace's operator

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}.$$

0.2 Some facts about integrals

Gauss formula. Consider in some finite-dimensional space, R_n , a quadratic form,

$$(Au, u) = A_{ij} u^i u^j, \quad (19)$$

The form is assumed to be positive,

$$(Au, u) > 0 \text{ if } u \neq 0.$$

Then the Gauss formula holds true:

$$\int_{R_n} e^{-\frac{1}{2}(Au, u)} du = \frac{1}{\sqrt{\det A}}. \quad (20)$$

Here

$$\det A \equiv \det \|A_{ij}\|, \quad du = \frac{du^1}{\sqrt{2\pi}} \dots \frac{du^n}{\sqrt{2\pi}}.$$

The Gauss formula can be proved by changing the variables, $u \rightarrow \hat{u}$,

$$u^i = \lambda_j^i \hat{u}^j, \quad \det \|\lambda_j^i\| = 1,$$

\hat{u}^j being the coordinates in which the tensor A_{ij} is diagonal,

$$(Au, u) = A_{ij} \lambda_j^i \lambda_{j'}^j \hat{u}^i \hat{u}^{j'} = A_1 (\hat{u}^1)^2 + \dots + A_n (\hat{u}^n)^2. \quad (21)$$

In the new variables,

$$\begin{aligned} \int_{R_n} e^{-\frac{1}{2}(Au, u)} du &= \int_{R_n} e^{-\frac{1}{2}(A_1 (\hat{u}^1)^2 + \dots + A_n (\hat{u}^n)^2)} d\hat{u} \\ &= \frac{1}{\sqrt{A_1 \dots A_n}} = \frac{1}{\sqrt{\det A}} \end{aligned}$$

Here we used that⁴

$$\int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2} dx = \sqrt{2\pi}. \quad (22)$$

The Gauss formula admits the following generalization: for any linear function of u , $(l, u) = l_i u^i$,

$$\sqrt{\det A} \int_{R_n} e^{-\frac{1}{2}(Au, u) + (l, u)} du = e^{\frac{1}{2}(A^{-1}l, l)} \quad (23)$$

where A^{-1} is the inverse matrix to the matrix A . Formula (23) follows from (20) and the identity,

$$\frac{1}{2} (Au, u) - (l, u) = \frac{1}{2} (A(u - A^{-1}l), (u - A^{-1}l)) - \frac{1}{2} (A^{-1}l, l). \quad (24)$$

Plugging (24) in (23), changing the variables of integration, $u \rightarrow u + A^{-1}l$, and using (20) we obtain the right hand side of (23).

⁴The integral (22) can be found by a witty trick suggested by Poisson:

$$\begin{aligned} \left(\int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2} dx \right)^2 &= \int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2} dx \int_{-\infty}^{+\infty} e^{-\frac{1}{2}y^2} dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2} e^{-\frac{1}{2}y^2} dx dy \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(x^2+y^2)} dx dy = \int_0^{+\infty} \int_0^{2\pi} e^{-\frac{1}{2}r^2} r dr d\theta = 2\pi \int_0^{+\infty} e^{-\frac{1}{2}r^2} d\frac{1}{2}r^2 = 2\pi. \end{aligned}$$

E x e r c i s e. Let (Au, u) be a positive quadratic form of variables $u = (u^1, \dots, u^n)$, and u be subject to s linear constraints,

$$(l_\alpha, u) = c_\alpha, \quad \alpha = 1, \dots, s; \quad s < n.$$

Consider an integral,

$$J = \int_{R_n} e^{-\frac{1}{2}(Au, u)} \prod_{\alpha=1}^s \delta(c_\alpha - (l_\alpha, u)) du^1 \dots du^n.$$

Show that

$$J = \sqrt{\frac{(2\pi)^n}{\det A}} \sqrt{\frac{(2\pi)^s}{\det \mathcal{A}}} e^{-\frac{1}{2}(\mathcal{A}^{-1}c, c)} \quad (25)$$

where \mathcal{A} is the matrix with components

$$\mathcal{A}_{\alpha\beta} = (A^{-1}l_\alpha, l_\beta).$$

H i n t. Use the presentation of δ -function (17) and formula (20).

Laplace's asymptotics. Consider an integral which depends on a parameter, λ , in the following way:

$$I(\lambda) = \int_V f(x) e^{\lambda S(x)} dV,$$

where V is a bounded region of n -dimensional space, $f(x)$ and $S(x)$ are some smooth functions. We wish to find the asymptotics of this integral as $\lambda \rightarrow \infty$. Laplace suggested that the leading terms of the asymptotics of $I(\lambda)$ are the same as that of the integral over the vicinities of the points where the function, $S(x)$, has the maximum value. Then the asymptotics can be easily found. Indeed, let $S(x)$ achieve its maximum value only at one point, \hat{x} , this point is an internal point of V , and the matrix of the second derivatives, $\|\partial^2 S(\hat{x})/\partial x^i \partial x^j\|$ is non-degenerated, i.e. its determinant, Δ , is non-zero. We can write,

$$I(\lambda) = f(\hat{x}) e^{\lambda S(\hat{x})} \int_V \frac{f(x)}{f(\hat{x})} e^{-\lambda[S(\hat{x}) - S(x)]} dV.$$

In a small vicinity of the point, \hat{x} , we can replace $S(\hat{x}) - S(x)$ by the non-degenerated quadratic form,

$$S(\hat{x}) - S(x) \approx -\frac{1}{2} S_{ij} (x^i - \hat{x}^i)(x^j - \hat{x}^j), \quad (26)$$

where $S_{ij} = \partial^2 S(\hat{x}) / \partial x^i \partial x^j$. Note that the quadratic form (26) is positive because \hat{x} is the point of maximum of $S(x)$. In a small vicinity of \hat{x} we can replace $f(x)/f(\hat{x})$ by unity thus obtaining

$$I(\lambda) \approx f(\hat{x}) e^{\lambda S(\hat{x})} \int_{\text{small vicinity of } \hat{x}} e^{-\lambda[-\frac{1}{2} S_{ij}(x^i - \hat{x}^i)(x^j - \hat{x}^j)]} dV. \quad (27)$$

Since $\lambda \rightarrow \infty$, the function, $\exp[-\lambda[-\frac{1}{2} S_{ij}(x^i - \hat{x}^i)(x^j - \hat{x}^j)]]$, decays very fast away from \hat{x} . We do not pause to justify that the expansion of the integration region from a small vicinity of \hat{x} to the entire space, R , causes only exponentially small corrections in (27). Thus, we can write:

$$I(\lambda) \approx f(\hat{x}) e^{\lambda S(\hat{x})} \int_R e^{-\lambda[-\frac{1}{2} S_{ij}(x^i - \hat{x}^i)(x^j - \hat{x}^j)]} dV.$$

The integral here, according to the Gauss formula (20), is equal to $\sqrt{(2\pi)^n / \lambda^n |\Delta|}$. Finally, the leading term of the asymptotics is

$$I(\lambda) \approx \sqrt{\frac{(2\pi)^n}{\lambda^n |\Delta|}} f(\hat{x}) e^{\lambda S(\hat{x})}. \quad (28)$$

As a more elaborated derivation shows, the error of the formula (28) is on the order of $(e^{\lambda S(\hat{x})} / \lambda^{n/2}) / \lambda$. If $S(x)$ achieves its maximum at several internal points, one should sum the contributions (28) of all points. One can check that in the cases of the point of maximum lying on the boundary and/or degeneration of the quadratic form $-\frac{1}{2} S_{ij}(x^i - \hat{x}^i)(x^j - \hat{x}^j)$ the asymptotics remains qualitatively the same,

$$I(\lambda) \approx \text{prefactor}(\lambda) e^{\lambda S(\hat{x})}, \quad (29)$$

with the prefactor being a decaying power function of λ .

The prefactor is a constant independent on λ , if $S(x)$ has maximum value on a set with non-zero volume. By Laplace's asymptotics we mean further the asymptotics of the form (29) where the prefactor changes slower than the exponential function of λ :

$$\frac{1}{\lambda} \ln \text{prefactor}(\lambda) \rightarrow 0 \quad \text{as } \lambda \rightarrow \infty.$$

Changing in the previous consideration $S(x)$ by $-S(x)$ we obtain the asymptotics,

$$\int_V f(x) e^{-\lambda S(x)} dV \approx \text{prefactor}(\lambda) e^{-\lambda S(\hat{x})}, \quad (30)$$

where \tilde{x} is the point of minimum of $S(x)$.

In applications to the variational problems, we need also to know the asymptotics of integrals of the form (30) for complex values of λ . In this case we denote the parameter by z ,

$$I(z) = \int_V f(x) e^{zS(x)} dV,$$

and consider the asymptotics of $I(z)$ as $|z| \rightarrow \infty$. Note first of all that $I(z)$ is an analytical function of z at any finite point, z , if the integral, as we accept, converges absolutely, i.e.

$$\int_V |f(x)| e^{\operatorname{Re} z S(x)} dV < \infty.$$

The point $z = \infty$ can be, however, the singular point of $I(z)$. Usually, the singularity is essential, i.e. the asymptotics of $I(z)$ along different paths, $z \rightarrow \infty$, are different. It turns out that Laplace's asymptotics,

$$I(z) \approx \sqrt{\frac{(2\pi)^n}{z^n |\Delta|}} f(\hat{x}) e^{zS(\hat{x})} \left(1 + O\left(\frac{1}{z}\right)\right), \quad (31)$$

holds true for all paths, $z \rightarrow \infty$, such that $|\operatorname{Arg} z| \leq \pi/2 - \varepsilon$, for some small ε . For other paths, this asymptotics does not hold. This is seen from studying the asymptotics when $z \rightarrow \infty$ along the imaginary axis, $z = iy$, $|y| \rightarrow \infty$. It turns out that in this case the leading contribution to the asymptotics is provided by not only the point of maximum of $S(x)$, but by all stationary points of $S(x)$, in particular, by all points of local maxima and minima. This asymptotics is called the stationary phase asymptotics; we do not dwell on it here since it will not be used further.

0.3 Reminder from probability theory

The random variables which we will be dealing with are the points, x , of some finite-dimensional region, V . Probability of the event that x belongs to a set A , $A \subset V$, is, by definition,

$$p(A) = \int_A f(x) dx.$$

Non-negative function $f(x)$ is called the probability density. Since $p(V) = 1$,

$$\int_V f(x) dx = 1.$$

Average value of function $\varphi(x)$ is defined as

$$M\varphi = \int_V \varphi(x)f(x)dx.$$

Here M stands for "mathematical expectation". This operation is often denoted by the symbol E (for expectation), but we reserve E to be the symbol of energy.

For any function φ , $\varphi(x)$ is a random variable. Probability density of $\varphi(x)$, f_φ , is defined in terms of probability of the event $\{\xi \leq \varphi(x) \leq \xi + \Delta\xi\}$ for small $\Delta\xi$:

$$f_\varphi(\xi)\Delta\xi = \text{Prob} \{\xi \leq \varphi(x) \leq \xi + \Delta\xi\}.$$

It is convenient to introduce the distribution function,

$$F(\xi) = \text{Prob} \{\varphi(x) \leq \xi\}.$$

If the distribution function is smooth, then

$$f_\varphi(\xi) = \frac{dF(\xi)}{d\xi}.$$

It is convenient to write the distribution function in terms of the step function

$$F(\xi) = M\theta(\xi - \varphi(x)) = \int \theta(\xi - \varphi(x)) f(x)dx. \quad (32)$$

This formula explains how the step function, to which we have paid already much attention, enters our consideration.

Consider a random variable, which is a couple (x, y) . Its probability density is a function of x and y , $f(x, y)$. Let y take values only in some region B . What would be a probability density of x ? It is natural to define probability density of x under condition that $y \in B$ as

$$f(x) = \int_B f(x, y) dy \Bigg/ \int_V \int_B f(x, y) dx dy.$$

Obviously,

$$\int f(x)dx = 1.$$

In general, the conditional probability density of x depends on the choice of B . By definition the random variables x and y are statistically independent,

if the conditional probability does not depend on B . This is possible only if $f(x, y)$ is the product of two functions

$$f(x, y) = f(x)g(y), \quad \int_V f(x)dx = 1, \quad \int_V g(y)dy = 1. \quad (33)$$

The notion of statistical independence is the central one in probability theory. All most important facts of probability theory are concerned with the sets of independent events. The major applications of probability theory are based to the possibility to identify the independent (or slightly dependent) events⁵.

We prepared everything to solve the central for applications problem of probability theory. Let x_1, \dots, x_n be independent identically distributed random variables. Find probability distribution of the sum

$$\varphi(x_1) + \dots + \varphi(x_N)$$

where φ is a given function. Of course, we expect to get an analytical answer only in the limit of large N .

Let us find the probability distribution of the arithmetic average

$$E = \frac{1}{N} (\varphi(x_1) + \dots + \varphi(x_N)).$$

We have

$$F_N(E) = M\theta\left(E - \frac{1}{N} (\varphi(x_1) + \dots + \varphi(x_N))\right).$$

We use for the sum the symbol E because in similar problems to be considered later, it has the meaning of energy. Using the presentation of the step function (10) we have

$$F_N(E) = M \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{dz}{z} e^{Ez - z \frac{1}{N} \varphi(x_1) - \dots - z \frac{1}{N} \varphi(x_N)}. \quad (34)$$

⁵By the way, the recent financial crisis was caused in part by a wrong identification of independent events. The hedge fund traders believed that combining mortgage loans into large packages for sale to banks and pension funds reduces the risk of default. This is true if the defaults of individual loans were independent. However, as we have experienced, this is not always the case: there are rare catastrophic events when probabilities of defaults become strongly correlated. This is what happened in the recent economic crisis: simultaneous default of many mortgage loans along with other negative events drove the economy down, which, in turn, resulted in more mortgage defaults and bankruptcy of financial institutions.

It is convenient to change variable z by zN . Then

$$F_N(E) = M \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{dz}{z} e^{NEz - z\varphi(x_1) - \dots - z\varphi(x_N)}.$$

In such change, the constant a in the integral limit must be replaced by Na , but, since this constant is arbitrary, we keep the some notations for the integral limit. The operation of mathematical expectation is, in essence, integration. The order of integrals can be changed almost always; we do not pay attention to degenerated cases. Since the variables x_1, \dots, x_N are statistically independent and identically distributed,

$$Me^{-z\varphi(x_1) - \dots - z\varphi(x_N)} = \left(Me^{-z\varphi(x)} \right)^N,$$

and we get

$$\begin{aligned} F_N(E) &= \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{dz}{z} e^{NEz} \left(Me^{-z\varphi(x)} \right)^N \\ &= \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{dz}{z} e^{NEz + N \ln Me^{-z\varphi(x)}} = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{dz}{z} e^{NS(E,z)} \end{aligned} \quad (35)$$

where we introduced a function of E and z

$$S(E, z) = Ez + \ln Q(z), \quad Q(z) \equiv \int e^{-z\varphi(x)} f(x) dx. \quad (36)$$

In physical applications, functions S has the meaning of entropy, and we will call S the entropy of the problem.

The integral (35) contains a large parameter, N . Therefore, probability distribution $F_N(E)$ for large N can be found by studying the asymptotics of the integral (35) as $N \rightarrow \infty$. This idea is in the core of all further examples considered.

Usually, the integral in (35) can be differentiated over E . After differentiation we obtain formula for probability density of the normalized sum,

$$f_N(E) = \frac{N}{2\pi i} \int_{a-i\infty}^{a+i\infty} e^{NS(E,z)} dz. \quad (37)$$

Note that the integrand, which is equal to

$$e^{NEz} \left(\int e^{-z\varphi(x)} f(x) dx \right)^N,$$

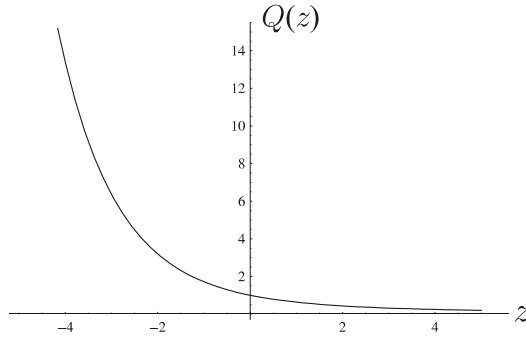


Figure 6. Plot of $Q(z)$.

does not have singularities for finite z , and, if it decays as $|Imz| \rightarrow \infty$, the line of integration can be moved to the left half-plane. Therefore, a in (37) can take both positive and negative values.

We consider examples in the next lecture.

0.4 The central limit theorem and the law of large numbers

We begin with the following example. Let x be a random number on the segment $[0,1]$, which is homogeneously distributed, i.e. $f(x) \equiv 1$. Then $Q(z)$ is computed analytically:

$$Q(z) = \int_0^1 e^{-zx} dx = \frac{1}{z} (1 - e^{-z}).$$

Graph of $Q(z)$ is shown in Fig. 6.

Function $S(E, z)$ is shown in Fig. 7.

Function $\ln Q(z)$ is a convex function of z (we will prove it in a more general case later). Therefore, $S = Ez + \ln Q$ is also convex and has a minimum. To find the minimizer we have to solve the equation

$$\frac{\partial S(E, z)}{\partial z} = E + \frac{Q'(z)}{Q(z)} = 0.$$

It can be written as

$$-\frac{Q'(z)}{Q(z)} = \frac{1}{z} - \frac{e^{-z}}{1 - e^{-z}} = E. \quad (38)$$

Function $-Q'(z)/Q(z)$ takes the values between 0 and 1 (again, we show this later in a general case). Therefore, for $0 \leq E \leq 1$, equation (38) has a

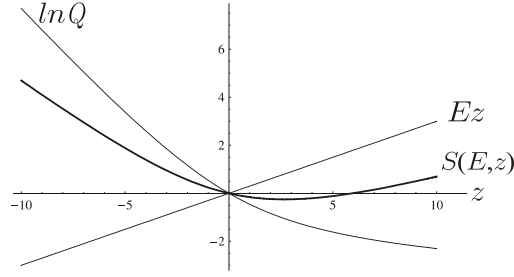


Figure 7. Plot of $\ln Q(z)$, Ez and $S(E, z)$ for $E = 0.3$.

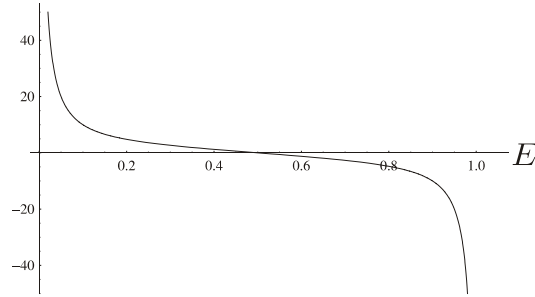


Figure 8. Dependence of inverse temperature β on energy E .

unique solution, which we denote $\beta(E)$. This notation is motivated again by physical problems: in similar physical problems β plays the role of inverse temperature. Function $\beta(E)$ is shown in Fig. 8.

If we tend z to zero in (38), we find

$$\frac{1}{z} - \frac{1 - z + \frac{1}{2}z^2}{1 - (1 - z + \frac{1}{2}z^2)} \rightarrow \frac{1}{2}.$$

Therefore, the value of E at which $\beta = 0$, is $1/2$. If $E < 0$, then $S(E, z)$ has the minimum at $z = +\infty$, if $E > 1$, $S(E, z)$ has the minimum at $z = -\infty$ (Fig. 9).

Denote by $S(E)$ the value of $S(E, z)$ at the point of minimum over z ,

$$S(E) = S(E, \beta(E)).$$

The graph of function $S(E)$ is shown in Fig. 10.

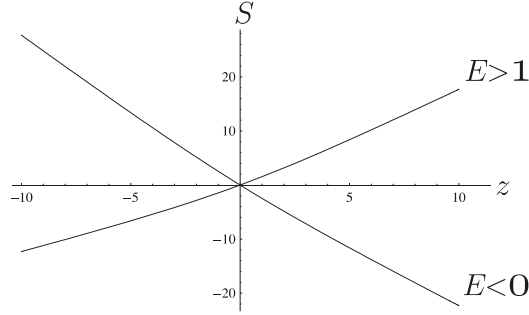


Figure 9. A qualitative graph of $S(E, z)$ as a function of z for $E > 1$ and $E < 0$.

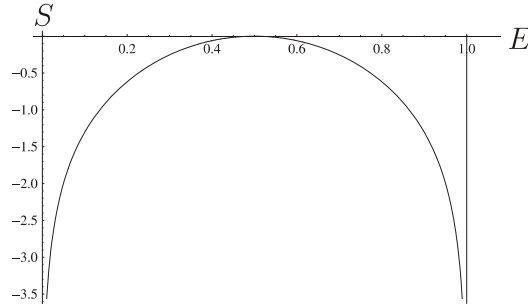


Figure 10. Dependence of $S(E)$ on E .

E x e r c i s e. Show that

$$\beta(E) = \frac{dS(E)}{dE}.$$

In the sum, $x_1 + \dots + x_N$, each member takes the values between 0 and 1. Therefore, the admissible values of the sum lie between 0 and N , while the values of E are between 0 and 1. Probability that $E < 0$ or $E > 1$ is zero. This fact can be derived directly from (35), (36), but we do not pause on this derivation and focus on the values of E from 0 to 1. We set a in (37) equal to $\beta(E)$. Formula (37) takes the form

$$f_N(E) = \frac{N}{2\pi} \int_{-\infty}^{+\infty} e^{NS(E, \beta + iy)} dy \quad (39)$$

$$S(E, z) = Ez + \ln \left[\frac{1 - e^{-z}}{z} \right].$$

Since $f_N(E)$ is real, the imaginary part of the integral (39) is zero.

Consider function $S(E, \beta + iy)$ for small y . Since $\partial S(E, z) / \partial z = 0$ at $y = 0$, we have

$$S(E, z) = S(E) + \frac{1}{2} S''(E) (iy)^2 = S(E) - \frac{1}{2} S''(E) y^2, \quad (40)$$

$$S''(E) \equiv \frac{\partial^2 S(E, z)}{\partial z^2} \Big|_{z=\beta(E)} = \frac{\partial^2 \ln Q(z)}{\partial z^2} \Big|_{z=\beta(E)}.$$

Function $S(E, z)$ is strictly convex on real axis, thus $S''(E) > 0$, and $S(E, \beta + iy)$ has a local maximum on y -axis at $y = 0$. For finite y , $\text{Re} S(E, \beta + iy)$ decays. Indeed,

$$\begin{aligned} \left| \ln \left[\frac{1 - e^{-\beta - iy}}{\beta + iy} \right] \right| &= \ln \frac{|1 - e^{-\beta - iy}|}{\sqrt{\beta^2 + y^2}} \\ &= \ln(e^{-\beta} |e^{\beta} - e^{-iy}|) - \ln \sqrt{\beta^2 + y^2}. \end{aligned}$$

The first term here is bounded, while the second one goes to $-\infty$. So, the major contribution to this integral as $N \rightarrow \infty$ is provided by a vicinity of the point $y = 0$. Replacing $S(E, z)$ by (40) we have

$$f_N(E) = e^{NS(E)} \frac{N}{2\pi} \int_{-\infty}^{+\infty} e^{-\frac{N}{2} S''(E) y^2} dy = \sqrt{\frac{N}{2\pi S''(E)}} e^{NS(E)}. \quad (41)$$

Here we made the change of variable $y \rightarrow y / \sqrt{NS''}$ and used (22).

Formula (41) is an asymptotic formula as $N \rightarrow \infty$. The normalization condition

$$\int_{-\infty}^{+\infty} f_N(E) dE = \int_0^1 f_N(E) dE = 1 \quad (42)$$

is satisfied asymptotically. Indeed, in the integral

$$\int_{-1}^1 \frac{1}{\sqrt{S''(E)}} e^{NS(E)} dE$$

the major contribution is provided by the vicinity of the point of maximum of $S(E)$, which is $\hat{E} = \frac{1}{2}$. At this point $S = 0$; thus

$$\int_{-1}^1 \frac{1}{\sqrt{S''(E)}} e^{NS(E)} dE \simeq \int_{-\infty}^{+\infty} \frac{1}{\sqrt{S''(E)}} e^{\frac{1}{2}NS_{EE}(\hat{E})(E-\hat{E})^2} dE. \quad (43)$$

Here we denoted by $S_{EE}(E)$ the second derivative of $S(E)$:

$$S_{EE}(E) = \frac{d^2 S(E)}{dE^2}.$$

The derivative $S_{EE}(E)$ is negative, because $S(E)$ is a concave function. For the integral (43) we have in the leading approximation,

$$\frac{\sqrt{2\pi/N}}{\sqrt{-S_{EE}(\hat{E})}} \frac{1}{\sqrt{S''(\hat{E})}}. \quad (44)$$

From (38) and (40)

$$S''(\hat{E}) = \left. \frac{d^2 S(E, z)}{dz^2} \right|_{z=\beta(\hat{E})} = \left. \frac{d^2 \ln Q}{dz^2} \right|_{z=\beta(\hat{E})} = - \left. \frac{dE}{d\beta} \right|_{z=\beta(\hat{E})}.$$

Hence

$$S''(\hat{E}) \cdot S_{EE}(\hat{E}) = S''(\hat{E}) \left. \frac{d\beta}{dE} \right|_{\hat{E}} = -1. \quad (45)$$

Combining (44), (45) and (41) we obtain (42).

Our asymptotic result converges to the exact one very fast. For $N = 10$, the exact and asymptotic results are shown in Fig. 11.

For $N = 100$ the exact and asymptotic results are practically indistinguishable.

As N increases, probability density converges to δ -function, concentrated at the point $E = \frac{1}{2}$. This value, $E = \frac{1}{2}$, is the mathematical expectation of each of the members of the sum. We obtain *the law of large numbers*:

$$\frac{1}{N} (x_1 + \dots + x_N) \rightarrow Mx = \frac{1}{2} \quad \text{as } N \rightarrow \infty. \quad (46)$$

Equation (46) can be interpreted in the following way. Let we do some experiments and the outcome of the experiment is a number, x , $0 \leq x \leq 1$. All values of x on the segment $[0, 1]$ are equiprobable. The outcomes of the experiments are independent. Then the arithmetic average of all outcomes for large N is approximately equal to the mathematical expectation of x .

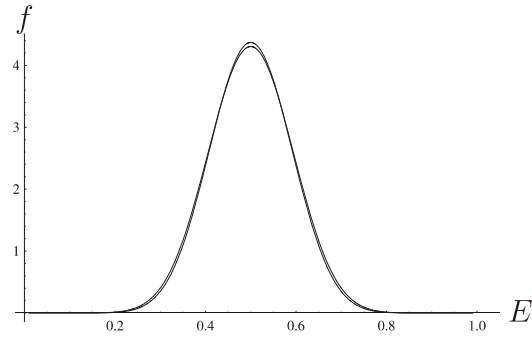


Figure 11. Exact probability density of E compared with the asymptotic formula when $N = 10$; the exact and asymptotic results are hardly distinguishable.

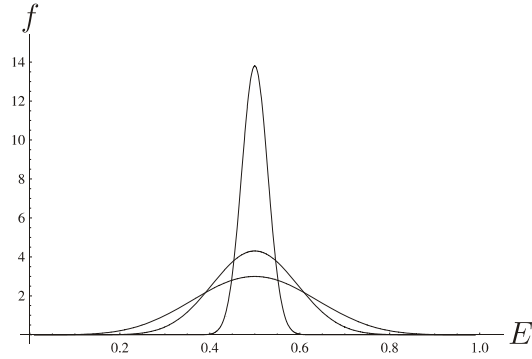


Figure 12. Exact probability densities of E for $N = 5, 10$ and 100 .

It is interesting to characterize the deviations of the arithmetic average from the mathematical expectation,

$$\frac{1}{N} (x_1 + \dots + x_N) - Mx = E'.$$

To estimate the order of deviations, let us find the mathematical expectation of E'^2 . We can do that using the probability density of E (41). From (41) we have

$$\begin{aligned} ME'^2 &= M \left(E - \frac{1}{2} \right)^2 = \int_0^1 \left(E - \frac{1}{2} \right)^2 \sqrt{\frac{N}{2\pi S''(E)}} e^{NS(E)} dE \\ &= \sqrt{\frac{N}{2\pi S''(\frac{1}{2})}} \int_{-\infty}^{+\infty} E'^2 e^{NS_{EE}(\frac{1}{2})E'^2} dE' = \frac{1}{N} \frac{S''(\frac{1}{2})}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{1}{2}x^2} dx. \end{aligned}$$

Here we used that $S(E)$ has maximum at $E = \frac{1}{2}$ and equal to zero at this point, besides, equation (45) was also employed.

We see that ME'^2 is of the order $1/N$. Hence, E' is of the order $1/\sqrt{N}$. This suggests that the scaled deviations,

$$\xi = \sqrt{N}E',$$

are of the order of unity and can have a non-singular probability distribution. Let us find it. Denote the probability density function of ξ by $g(\xi)$. Since,

$$E = \frac{1}{2} + \frac{\xi}{\sqrt{N}}$$

and

$$f_N(E) dE = g(\xi) d\xi,$$

we have

$$g(\xi) = \frac{1}{\sqrt{N}} f_N \left(\frac{1}{2} + \frac{\xi}{\sqrt{N}} \right).$$

Plugging here (41) we find for large N

$$g(\xi) = \frac{1}{\sqrt{2\pi S''(\frac{1}{2})}} e^{\frac{1}{2} S_{EE}(\frac{1}{2}) \xi^2}.$$

Denoting $S''(\frac{1}{2})$ by σ^2 and using (45) we obtain

$$g(\xi) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma^2} \xi^2}. \quad (47)$$

This is Gaussian distribution. The constant σ has the meaning of variance of ξ ,

$$M\xi^2 = \sigma^2.$$

Formula (47) expresses the so-called *central limit theorem*. The law of large numbers and the central limit theorem are simple consequences of (41). Formula (41) provides much more information: it determines the probability of large deviations of the arithmetical average from the average value. This probability is exponentially small because $S(E)$ is negative.

E x e r c i s e s. 1. Let u_a be non-negative numbers, $a = 1, \dots, N$. All points of the space $\{u_1, \dots, u_N\}$ are equiprobable. Note that we cannot introduce probability in a usual sense because the volume of the admissible values is infinite. Let u_a be constrained by the condition

$$\frac{1}{N} \sum_{a=1}^N u_a = 1.$$

This condition makes the volume of the admissible values finite. Show that in the limit $N \rightarrow \infty$ the values of any two numbers (say, u_1 and u_2) become statistically independent, and each number has the probability distribution

$$f(u) = e^{-u}.$$

2. Let a_1, \dots, a_n, \dots be an infinite sequence of numbers and x_1, \dots, x_n, \dots a sequence of independent identically distributed variables with mathematical expectation Mx , variance Mx^2 and probability density function $f(x)$. Consider a random number

$$\xi = a_1x_1 + \dots + a_nx_n + \dots$$

show that

$$M\xi = Mx \sum_{k=1}^{\infty} a_k$$

$$M\xi^2 = Mx^2 \sum_{k=1}^{\infty} a_k^2$$

while the probability density function of ξ , $f_\xi(y)$ is given by the integral

$$f_\xi(y) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} e^{zy-g(y)} dz,$$

$$g(y) = \sum_{k=1}^{\infty} \ln \int e^{-y a_k x} f(x) dx.$$

It is assumed that the sums and integrals in these formulas exist. A necessary condition for convergence of series is the decay a_k as $k \rightarrow \infty$.

0.5 Poisson distribution

Now we return to a more general case we have started with, when x was a point of some region distributed with some probability density $f(x)$. As we will see all results we obtained for numbers are extended to this case. We have obtained already the formulas for probability density (37), (36). Consider the properties of $S(E, z)$ on real axis. This function is a convex function of z . Indeed,

$$\frac{\partial S(E, z)}{\partial z} = E - \frac{\int \varphi e^{-z\varphi} f dx}{\int e^{-z\varphi} f dx} \quad (48)$$

$$\frac{\partial^2 S(E, z)}{\partial z^2} = \frac{\int \varphi^2 e^{-z\varphi} f dx \int e^{-z\varphi} f dx - \left(\int \varphi e^{-z\varphi} f dx \right)^2}{\left(\int e^{-z\varphi} f dx \right)^2}. \quad (49)$$

Using Cauchy inequality,

$$\left(\int f \cdot g dx \right)^2 \leq \int f^2 dx \int g^2 dx,$$

we have

$$\begin{aligned} \left(\int \varphi e^{-z\varphi} f dx \right)^2 &= \left(\int \varphi e^{-\frac{1}{2}z\varphi} f^{\frac{1}{2}} \cdot e^{-\frac{1}{2}z\varphi} f^{\frac{1}{2}} dx \right)^2 \\ &\leq \int \varphi^2 e^{-z\varphi} f dx \int e^{-z\varphi} f dx. \end{aligned}$$

Therefore,

$$\frac{\partial^2 S(E, z)}{\partial z^2} \geq 0,$$

and $S(E, z)$ is a convex function of z . Hence, it may have only one local minimum at a finite z . It may have also minimum at $z = +\infty$ or $z = -\infty$. Consider the case when minimum is achieved at a finite point, \tilde{z} . According to (48), \tilde{z} is the solution of the equation

$$\frac{\int \varphi e^{-z\varphi} f dx}{\int e^{-z\varphi} f dx} = E. \quad (50)$$

We assume that function $\varphi(x)$ is piecewise continuous and bounded in V and has the minimum and maximum values, φ_- and φ_+ . Since

$$\varphi_- \int e^{-z\varphi} f dx \leq \int \varphi e^{-z\varphi} f dx \leq \varphi_+ \int e^{-z\varphi} f dx,$$

the left hand side of (50) is within the limits

$$\varphi_- \leq \frac{\int \varphi e^{-z\varphi} f dx}{\int e^{-z\varphi} f dx} \leq \varphi_+.$$

Therefore, the solution of equation (50) exists only for the values of E belonging to the segment

$$\varphi_- \leq E \leq \varphi_+.$$

According to Laplace asymptotics, the left hand side of (50) tends to φ_- as $z \rightarrow +\infty$ and φ_+ as $z \rightarrow -\infty$. So, the picture is completely similar to that of the case of random numbers.

Consider one special case which has a lot of applications. Let us choose $\varphi(x)$ to be a characteristic function of some subregion B of volume V , i.e.

$$\varphi(x) = \begin{cases} 1 & \text{if } x \in B \\ 0 & \text{otherwise.} \end{cases}$$

Then the sum,

$$\mathcal{N} = \varphi(x_1) + \dots + \varphi(x_N),$$

has the meaning of the number of points which are in the region B . This number is random and takes the values $0, 1, \dots, N$. We wish to find the probabilities that \mathcal{N} has values $0, 1, \dots$. To this end, we have to rewrite formula (35) in terms of probability distribution of the non-scaled sum, $\mathcal{N} = NE$. Denotes its values by k . Repeating the derivation from (34) to (35) we have

$$\begin{aligned} F_N(k) &= M \frac{1}{2\pi i} \int_{a-\infty}^{a+\infty} \frac{dz}{z} e^{kz - z\varphi(x_1) - \dots - z\varphi(x_N)} \\ &= \frac{1}{2\pi i} \int_{a-\infty}^{a+\infty} \frac{dz}{z} e^{kz} \left(\int e^{-z\varphi(x)} f(x) dx \right)^N. \end{aligned} \quad (51)$$

Let all points be homogeneously distributed over V , i.e.

$$f(x) = \frac{1}{|V|} = \text{const.}$$

Then

$$\int e^{-z\varphi(x)} f(x) dx = \frac{1}{|V|} (|V| - |B| + |B| e^{-z}). \quad (52)$$

Note that region B is not necessarily simply connected and may consist of many pieces, but formula (52) contains only the volume of region B . We consider the "thermodynamic limit", when $|V| \rightarrow \infty$, $N \rightarrow \infty$, while the number of points per unit volume,

$$n = \frac{N}{|V|},$$

remains finite. We have

$$\left(\int e^{-z\varphi(x)} f(x) dx \right)^N = \left(1 - \frac{|B|n}{N} (1 - e^{-z}) \right)^N.$$

In the limit $N \rightarrow \infty$

$$\left(\int e^{-z\varphi(x)} f(x) dx \right)^N = e^{-|B|n(1-e^{-z})}.$$

Hence, (51) takes the form

$$F_\infty(k) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{dz}{z} e^{kz - |B|n(1-e^{-z})}. \quad (53)$$

We replace N by ∞ in notation of distribution function because the right hand side of (53) is the limit as $N \rightarrow \infty$. One can show that (53) can be differentiated over k . We get for probability density

$$\frac{dF_\infty(k)}{dk} = f_\infty(k) = e^{-|B|n} \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} e^{kz + |B|ne^{-z}} dz.$$

This integral can be computed analytically. Indeed, let us present $\exp(|B|ne^{-z})$ as the series

$$e^{|B|ne^{-z}} = 1 + |B|ne^{-z} + \frac{1}{2}(|B|n)^2 e^{-2z} + \dots \frac{1}{s!} (|B|n)^s e^{-sz} + \dots$$

According to (17),

$$\frac{1}{2\pi i} \int e^{kz} \frac{1}{s!} (|B|n)^s e^{-sz} dz = \delta(k-s) \frac{1}{s!} (|B|n)^s.$$

So,

$$f_{\infty}(k) = e^{-|B|n} \left(\delta(k) + |B|n \delta(k-1) + \frac{1}{2} (|B|n)^2 \delta(k-2) + \dots \right),$$

i.e. $f_{\infty}(k)$ is a sum of δ -functions, concentrated at the points $k = 0, 1, 2, \dots$. This means that the region B contains k points with probability

$$p_k = \frac{1}{k!} (|B|n)^k e^{-|B|n}.$$

The sum of all probabilities is equal to 1,

$$\sum_0^{\infty} p_k = 1,$$

as follows from Taylor expansion of exponential function

$$e^{|B|n} = \sum_{k=0}^{\infty} \frac{1}{k!} (|B|n)^k.$$

We arrived at the so-called *Poisson distribution*.

0.6 Stochastic variational problems

Many problems of physics and mechanics can be formulated as variational problems, i.e. as problems of minimization of some functional, $I(u)$, on a set of elements, u . We will consider the simplest case, when $I(u)$ is a quadratic functional, i.e. the functional of the form

$$I(u) = \frac{1}{2} (Au, u) - (l, u). \quad (54)$$

By u one can mean a point of a multidimensional space, $u = (u^1, \dots, u^n)$, (l, u) a linear function

$$(l, u) = l_i u^i, \quad (55)$$

and (Au, u) a quadratic function

$$(Au, u) = A_{ij} u^i u^j. \quad (56)$$

In continuum mechanics problems, one considers the limit $n \rightarrow \infty$, but a finite-dimensional truncation of continuum mechanics problems (for example, by the finite-element method) returns us to the finite-dimensional case

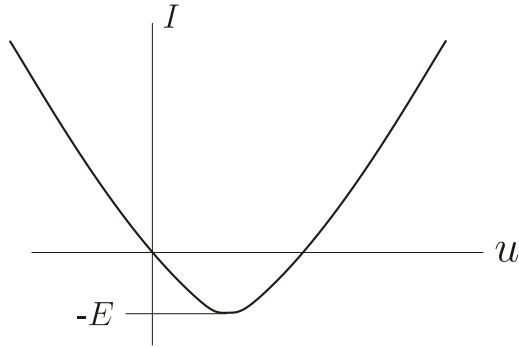


Figure 13. Minimum value of $I(u)$ is negative; it has the meaning of negative energy.

(55), (56). The minimum value of $I(u)$ is always negative; in 1D case this is seen from Fig. 13.

In physical problems the minimum value has the meaning of negative energy of the system.

If the properties of the system are random, so are the matrix (operator) $A = (A_{ij})$ and the vector (linear functional) $l = (l_i)$. We consider the simplest "probabilistic" question: What is the probability distribution of minimum values of $I(u)$ (i.e. probability distribution of energy)? It is enough to discuss the finite-dimensional case; the results for continuum mechanics are obtained in the limit $n \rightarrow \infty$.

If the matrix A does not depend on the event, ω , the problem is called weakly stochastic, otherwise the variational problem is called strongly stochastic. Many physical theories provide examples of such type of problems. The analytical results can be obtained mostly for weakly stochastic problems.

As is usual in the probabilistic approach, the probabilistic modeling is especially effective, if one can identify in the phenomenon to be modeled the statistically independent (or slightly dependent) events. Analytical investigation can be advanced considerably, if there are many statistically independent events involved. We focus here on a special case when (l, u) is a sum of small independent linear functionals. More precisely, there is a large number, N , of independent identically distributed random variables, r_1, \dots, r_N , and a given random linear functional, $(l_0(r), u)$. Then the linear functional of the variational problem, (l, u) , is defined as an "empirical average" of N values of $(l_0(r), u)$:

$$(l, u) = \frac{1}{N} \sum_{a=1}^N (l_0(r_a), u). \quad (57)$$

Consider a quadratic function of a finite number of variables (54). The minimum value of this function is

$$\tilde{I} = -\frac{1}{2} (A^{-1}l, l) \quad (58)$$

where A^{-1} is the inverse matrix for the matrix A . We wish to find probability distribution of energy,

$$f(E) = M \delta \left(E + \min_u I(u) \right). \quad (59)$$

Following the same path as for a sum of independent random variables we plug in (59) the presentation of δ -function (17),

$$f(E) = M \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} e^{Ez + z \min_u I(u)} dz = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} e^{Ez} M e^{z \min_u I(u)} dz. \quad (60)$$

It would be desirable to change somehow the operations of mathematical expectation and minimization: then we would arrive to some deterministic problem. This can be achieved by presenting $\exp[-z \min I(u)]$ by an integral of some function of u over u . Since mathematical expectation is, in fact, also integration, the order of integrals can be changed, and we obtain an integral of the mathematical expectation of the function of u , which can be found explicitly in some cases. Now let us discuss precisely what this trick means.

According to (58) formula (23) can be also written as

$$e^{-\min_u I(u)} = \sqrt{\det A} \int e^{-I(u)} du \quad (61)$$

We see that this relation reduces the computation of the minimum value to integration indeed. Since any quadratic functional in variational problems of continuum mechanics admits a finite-dimensional truncation, one can write formula (61) for a finite-dimensional truncation, and then consider the limit when the dimension of the truncation tends to infinity. In the limit, in the right hand side of (61) we obtain what is called the functional integral. We include $\sqrt{\det A}$ in the definition of the “volume element” in the functional space,

$$\mathcal{D}_A u = \sqrt{\det A} du \quad (62)$$

and write (61) as

$$e^{\min_u I(u)} = \int e^{-I(u)} \mathcal{D}_A u. \quad (63)$$

The notation, $\mathcal{D}_A u$, emphasizes that the volume element depends on the operator A .

E x e r c i s e. Consider a variational problem with a set of constraints: minimize a quadratic functional,

$$I(u) = \frac{1}{2}(Au, u)$$

on all u obeying to linear constraints

$$(l_\alpha, u) = c_\alpha, \quad \alpha = 1, \dots, s.$$

Show that

$$e^{-\min I(u)} = \int e^{-I(u)} \prod_{\alpha} \delta(c_\alpha - (l_\alpha, u)) \mathcal{D}u$$

where

$$\mathcal{D}u = \sqrt{\frac{(2\pi)^n}{\det A}} \sqrt{\frac{(2\pi)^s}{\det \mathcal{A}}} du^1 \dots du^n$$

and \mathcal{A} is a matrix with the components

$$\mathcal{A}_{\alpha\beta} = (A^{-1}l_\alpha, l_\beta).$$

H i n t. Use (25).

In Section 5.12 of the above-cited book *Variational Principles of Continuum Mechanics* one can find various generalizations of (63) involving non-positive quadratic functionals and complex-valued functionals. We illustrate the idea using one of such generalizations,

$$e^{z \min_u I(u)} = \int_{-i\infty}^{i\infty} e^{z[\frac{1}{2}(Au, u) - (l, u)]} \mathcal{D}_{zA} u, \quad \text{for } \operatorname{Re} z > 0. \quad (64)$$

In (64) the parameter, z , is also included in the volume element: for m -dimensional truncation, $\mathcal{D}_{zA} u = \sqrt{z^m \det A} du$.

If we plug in (64) the linear functional (57), we get

$$M e^{z \min_u I(u)} = \int_{-i\infty}^{i\infty} M e^{z[\frac{1}{2}(Au, u) - (l, u)]} \mathcal{D}_{zA} u$$

$$\begin{aligned}
&= \int_{-i\infty}^{i\infty} e^{z\frac{1}{2}(Au,u)} M e^{-z\frac{1}{N}\sum_{a=1}^N(l_0(r_a),u)} \mathcal{D}_{zA}u \\
&= \int_{-i\infty}^{i\infty} e^{z\frac{1}{2}(Au,u)} \left(M e^{-z\frac{1}{N}(l_0(r),u)} \right)^N \mathcal{D}_{zA}u.
\end{aligned}$$

If we change variable $z \rightarrow Nz$, then probability density $f(E)$ (60) takes the form

$$f(E) = \frac{N}{2\pi i} \int_{-i\infty}^{i\infty} e^{NS(E,z,u)} \mathcal{D}_{NzA}u dz, \quad (65)$$

where $S(E, z, u)$ is a function that is independent on N ,

$$S(E, z, u) = Ez + \frac{z}{2}(Au, u) + \ln M e^{-z(l_0, u)}. \quad (66)$$

The functional integral (65) depends on a large parameter N . In many cases the asymptotics of this integral can be studied. It is determined by the stationary points of the entropy functional (66). Examples and further details can be found in the above-cited book.