# Math480/540, TOPICS IN MODERN MATH. What are numbers? Part II

A. A. Kirillov*

Jan. 2008

In a role of numbers can occur not only elements of a field or a skew-field. In this part I talk about other mathematical objects which are used as numbers. It is interesting that all of them had appeared first in "pure" mathematics and later were used in mathematical physics and some other applications. It support the thesis that there is only one mathematics in which some domains have already found application and other not yet.

## 1 Matrices as numbers

> Examiner: What is a multiple root of a polynomial?
> University entrant: Well, it is when we substitute a number in the polynomial and get zero. Then do it again and again get zero and so $k$ times... But on the $(k+1)$-st time the zero does not appear.
>
> From the mathematical folklore of Moscow State University

### 1.1 Matrices; basic facts

Square matrices give us the remarkable generalization of the notion of a number. Let us denote by $\mathrm{Mat}_n(K)$ the set of all $n \times n$ matrices with

*Department of Mathematics, The University of Pennsylvania, Philadelphia, PA 19104-6395 *E-mail address: kirillov@math.upenn.edu*

entries from a field $K$. These entries are usually denoted by $A_{i,j}$ (or by $a_{i,j}$) where $i$ is a row number and $j$ is the column number. The set $\mathrm{Mat}_n(K)$ form a $K$-algebra. It means that the elements of $\mathrm{Mat}_n(K)$ can be added, multiplied and mu tiplied by numbers (i.e., elements of $K$). Namely,the addition, subtraction and multiplication by a number are defined element-wise:

$$(A \pm B)_{i,j} = A_{i,j} + B_{i,j}, \qquad (\lambda \cdot A)_{i,j} = \lambda \cdot A_{i,j}$$

and multiplication is defined by the formula

$$(A \cdot B)_{i,j} = \sum_k A_{i,k} \cdot B_{k,j}. \tag{1}$$

This rule seems to be rather artificial. Sometimes, people ask: isn't it better (and more natural) to define the multiplication element-wise, as all other operations? The answer is no! The definition (1) defines much more interesting and important algebra, than the direct sum of $n^2$ copies of $K$. The point is that (1) follows from the geometric interpretation of a matrix $A \in \mathrm{Mat}_n(K)$ as a linear transformation of a $n$-dimensional vector space $V$ over $K$.[1]

Indeed, if $\{v_i\}$, $1 \leq i \leq n$, are the coordinates of a vector $v \in V$, then a matrix $A$ defines the transformation $v \to AV$, where

$$(Av)_i = \sum_j A_{i,j} v_j \tag{2}$$

and a direct computation shows that the product (i.e composition) of transformations (2) corresponds to the product of matrices by the rule (1).

Thus, we can add, subtract and multiply matrices like numbers. But there are are also distinctions. First, the multiplication of matrices is, in general not commutative: $AB \neq BA$. Second, a division by a matrix (from the right or from the left) is possible only if the matrix is invertible. It is known that $A$ is invertible iff $\det A \neq 0$.

**Exercise 1** *Let $\mathbb{F}_q$ be a field with $q$ elements. How many matrices are in $\mathrm{Mat}_n(\mathbb{F}_q)$ and how many of them are invertible?*

[1]Note that it is not a sole way to give a geometric sense to a matrix. One can, for instance, consider a matrix $A \in \mathrm{Mat}_n(K)$ as a bilinear form on $V$, i.e. a map from $V \times V$ to $K$, which is linear in both arguments. The difference can be seen when we change a basis in $V$. The matrix of a linear operator is transformed according to the rule $A \to Q^{-1}AQ$, and a matrix of a bilinear form – according to another rule $A \to Q^t AQ$. Here $Q^{-1}$ is the inverse matrix and $Q^t$ is the transposed matrix.

In this section a matrix $A$ always corresponds to a linear operator.

The lack of commutativity prevented for a long time the use of matrices in a role of numbers. Only after appearance of quantum mechanics the "non-commutative numbers" became popular. We discuss it later and now consider another important feature of numbers. Namely, numbers can serve as arguments in functions. Let us look, how far matrices can be used in this capacity.

## 1.2 Matrix polynomials

The simplest functions studied in analysis are polynomials. Let $P(x) = \sum_{k=0}^{m} c_k x^k$ be a polynomial with real coefficients. It is clear, how to give sense to an expression of the form $P(A)$ where $A \in \mathrm{Mat}_n(\mathbb{R})$. We have to put

$$P(A) := \sum_{k=0}^{m} c_k A^k \qquad (3)$$

where, by definition, $A^0$ is a unit matrix which we denote by 1 (or by $1_n$, if it is necessary to indicate the size of a matrix).

**Exercise 2** *Let $\mathcal{P} := \mathbb{R}[x]$ be the polynomial algebra over $\mathbb{R}$ with one generator $x$. Find all homomorphisms $\varphi$ of $\mathcal{P}$ (in the category of associative real algebras with a unit) to the algebra $\mathrm{Mat}_n(\mathbb{R})$.*

One of the oldest problems of in algebra: to find the roots of a given polynomial $P(x)$. Later it turns out that the inverse problem is also of interest: to describe all polynomials for which a given number is a root. The answer is given by the

**Theorem 1 (Bezout Theorem)** *The set of all polynomials for which a given number $a$ is a root is an ideal in $\mathcal{P}$ generated by $x - a$.*[2]

Let us look, which restrictions on polynomial $P$ are imposed by the equation

$$P(A) = 0. \qquad (4)$$

The answer is rather simple, if the matrix $A$ is diagonal:

$$A = \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_n \end{pmatrix} = \mathrm{diag}(a_1,\, a_2,\, \dots,\, a_n).$$

[2] A more habitual formulation: $P(a) = 0$ iff $P$ is divisible by $(x - a)$.

In this case we have $P(A) = \operatorname{diag}\big(P(a_1),\, P(a_2),\, \ldots,\, P(a_n)\big)$. Therefore (4) is equivalent to the system

$$P(a_i) = 0, \qquad 1 \le i \le n. \tag{5}$$

So, the sole matrix $A$ replaces the whole set $\{a_1,\, a_2,\, \ldots,\, a_n\}$.

Let now the matrix $A$ be not necessarily diagonal, but can be reduced to the diagonal form:

$$A = Q^{-1}\Lambda Q \quad \text{where} \quad \Lambda = \operatorname{diag}(\lambda_1,\, \lambda_2,\, \ldots,\, \lambda_n).$$

It easy to check (but much more important to understand) that

$$P(Q^{-1}AQ) = Q^{-1}P(A)Q. \tag{6}$$

Therefore the equation (4) is equivalent to the system $P(\lambda_i) = 0,\ 1 \le i \le n$.

We see again that a matrix $A \in \operatorname{Mat}_n(\mathbb{R})$ replaces a set of $n$ numbers. From the linear algebra course you know that these numbers are *eigenvalues* of the matrix $A$, i.e. the roots of the *characteristic equation*

$$\det(A - \lambda \cdot 1) = 0.$$

The eigenvalues form a set $S$ which is called *spectrum* of $A$. If all the eigenvalues are different, we say that $A$ has a *simple spectrum.* It is known that any matrix with a simple spectrum in $K$ can be diagonalized over $K$. We come to the theorem.

**Theorem 2** *Let $A$ be a matrix with a simple spectrum $S$ which is contained in $K$. Then a polynomial $P \in K[x]$ vanishes at $A$ iff $P|_S = 0$.*

**Exercise 3** *Show that the theorem remains true even when the spectrum is not in $K$ but in some extension $\widetilde{K} \supset K$.*

The theorem is one of manifestation of the general principle:

*The matrix elements form only a perishable body of the operator A, while the eigenvalues express its immortal soul.*

Other manifestations of this principle occur below. Now I suggest the following subject for meditation.

**Problem.** Let $[A]$ denote a point of the projective space $\mathbb{P}^{n^2-1}(\mathbb{R})$, corresponding to a non-zero matrix $A \in \operatorname{Mat}_m(\mathbb{R})$.

Describe the behavior of the sequence $\{[A^n]\}$ (in particular, find out when it has a limit and how the set of limit points can look like).

### 1.3 Matrices and field extensions

For a given matrix $A \in \mathrm{Mat}_n(K)$ we consider the set $K[A]$ of all matrices of the form $P(A)$, $P \in K[x]$. It is clear that $K[A]$ is the minimal subalgebra with unit in $\mathrm{Mat}_n(K)$ which contains $A$. On the other hand, $K[A]$ is isomorphic to a quotient of $K[x]$ by the ideal $I(A)$ defined by (4).

**Theorem 3** *If $A$ is has a simple spectrum $S \subset K$, then the algebra $K[A]$ is isomorphic to the algebra $K[S]$ of $K$-valued functions on $S$.*

Indeed, from (6) it follows that the isomorphism in question can be defined as follows:

$$K[A] \ni P(A) \mapsto \big(P(\lambda_1), \ldots, P(\lambda_n)\big) \in K(S).$$

More interesting situation arises when the spectrum of $A$ is not simple or is not contained in $K$.

The model example of a matrix with non-simple spectrum is the so-called *Jordan block*:

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & \lambda \end{pmatrix}$$

**Exercise 4** *Prove that for $A = J_n(\lambda)$ the ideal $I(A)$ consists of polynomials $P$ satisfying*

$$P(\lambda) = P'(\lambda) = \ \ldots \ P^{(n-1)}(\lambda) = 0. \tag{7}$$

As is well-known (and easy to check), the spectrum of $J_n(\lambda)$ consists of a single point $\lambda$. It shows that the analogue of the theorems (2) and (3) are wrong.

However, from an intuitive point of view, it is convenient to believe that these theorem are still valid, even if for this we have to change some previous definitions.[3] In our case we have to change the definition of a spectrum and the class of functions considered. We want to consider the spectrum of $J_n(\lambda)$ not as a single point $\lambda$, but as an *infinitesimal neighborhood* of order $n-1$ of this point which we denote $U_{n-1}(\lambda)$.

[3] Here, as in many other cases, one can use the principle:

If a definition impedes the validity of a nice theorem, change it.

The Theorem (3) suggests what how we must understand the restriction of a polynomial $P$ to $U_{n-1}(\lambda)$: we have to develop $P$ in a power series in $(x-\lambda)$ and cut it after term of degree $n-1$. Then conditions (7) will be equivalent to $P|_{U_{n-1}(\lambda)} = 0$. In case $K = \mathbb{R}$ this construction is well-known in analysis and can be applied to any smooth function defined in an ordinary neighborhood of $\lambda$. More precisely, a $n$-jet of a function $f$ at the point $\lambda$ is the expression

$$j_n(f) = \sum_{k=0}^{n} f^{(k)}(\lambda) \cdot \frac{(x-\lambda)^k}{k!},$$

which we consider as the restriction of $f$ to $U_n(\lambda)$.

Of course, we can not consider $U_n(\lambda)$ as an ordinary neighborhood (compare $ 3). For example, the polynomial $P(x) = (x-\lambda)^n$ vanishes on $U_n(\lambda)$ but has a unique root $x = \lambda$.

Let us consider the case when the spectrum of $A$ is not contained in $K$. Then the characteristic polynomial

$$\chi_A(x) := \det(A - x \cdot 1) \tag{8}$$

does not decompose into linear factors. Let, for instance, $\chi_A$ is irreducible, i.e. does not decompose at all into factors of smaller degrees. Then the ideal $I(A) \subset K[x]$ is simple and generated by $\chi_A$. Therefore, the algebra $K(A) \simeq K[x]/I(A)$ has no zero divisors and is a field containing $K$. We denote it by $\widetilde{K}$.

**Lemma 1** *Let the matrix $A$ is of size $n \times n$, so that $\chi_A$ has degree $n$. Then, the field $\widetilde{K}$ as a vector space over $K$, has dimension $n$.*

Indeed, the elements $1, A, A^2, \ldots, A^{n-1}$ are linearly independent, because $I(A)$ contains no polynomials of degree $\leq n-1$. On the other hand, from the Cayley identity

$$\chi_A(A) = 0 \tag{9}$$

every element of $\widetilde{K}$ can be written in the form

$$a_0 \cdot 1 + a_1 \cdot A + \ldots + a_{n-1} \cdot A^{n-1}, \tag{10}$$

where $a_i \in K$.

If we want the theorem (3) to be still valid, there are two ways:

1. We can think that Spec A, the spectrum of $A$, consists of one point $\widetilde{\lambda}$, not belonging to $K$, and $K[A] \simeq \widetilde{K}$ is interpreted as the set of all functions on this single point $\widetilde{\lambda}$, taking values in $\widetilde{K}$.

2. We can think that Spec A consists of all $n$ eigenvalues of $A$ in $\widetilde{K}$ and the algebra $K[A]$ consists of $\widetilde{K}$-valued functions on the spectrum which satisfy the additional condition. Namely, denote by $G = Gal(\widetilde{K}/K)$ the so-called *Galois group* of the field $\widetilde{K}$ over $K$. It consists of all automorphisms of the field $\widetilde{K}$ which fix all elements of $K$. The condition in question is

$$f(g \cdot \lambda) = g \cdot f(\lambda) \quad \text{for all} \quad g \in G,\, \lambda \in \text{Spec A}.$$

**Example 1** *If* $K = \mathbb{R}$, $A = \left\| \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} \right\|$, *then* $I(A)$ *is generated by the irreducible over* $\mathbb{R}$ *polynomial* $\chi_A(x) = x^2+1$. *Here the field* $\widetilde{K}$ *is isomorphic to* $\mathbb{C}$. *We get a well-known realization of* $\mathbb{C}$ *by real matrices of the form*

$$a_0 \cdot 1 + a_1 \cdot A = \left\| \begin{matrix} a_0 & a_1 \\ -a_1 & a_0 \end{matrix} \right\|.$$

**Exercise 5** *Find the group* $Gal(\mathbb{C}/\mathbb{R}$.

**Example 2** *If* $K = \mathbb{F}_2$, $A = \left\| \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} \right\|$, *then* $I(A)$ *generated by the irreducible over* $\mathbb{R}$ *polynomial* $\chi_A(x) = x^2+x+1$ *and the field* $\widetilde{K}$ *is isomorphic to* $\mathbb{F}_4$. *We obtain a simplest realization of the field* $\mathbb{F}_4$[4] *by matrices over* $\mathbb{F}_2$*:*

$$0 = \left\| \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \right\|, \quad 1 = \left\| \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \right\|, \quad x = \left\| \begin{matrix} 0 & 1 \\ 1 & 1 \end{matrix} \right\|, \quad x+1 = \left\| \begin{matrix} 1 & 1 \\ 1 & 0 \end{matrix} \right\|.$$

**Exercise 6** *Find the group* $Gal(\mathbb{F}_4/\mathbb{F}_2)$.

## 1.4 Operational calculus

We know now, more or less, what are polynomial functions of matrices. Let us look, how to define more general functions of a matrix argument. For this we turn from algebra to analysis and use the limit operation. It is well-known that many functions can be approximated by polynomials in one or another sense. So, an analytic function on $\mathbb{R}$ is approximated by partial sums of its Taylor series and this approximation is uniform on any closed segment $[a, b]$; any continuous function on a segment $[a, b] \in \mathbb{R}$ can be uniformly approximated by a sequence of polynomials; a continuous function on the

[4] Which, certainly, is different from the ring $\mathbb{Z}_4 = \mathbb{Z}/4\mathbb{Z}$.

circle $x^2 + y^2 = 1$ is uniformly approximated by polynomials $P(x, y)$, or trigonometric polynomials $\sum_{k=-N}^{N} a_k \cos kx + b_k \sin kx$.

Let us choose a vector space $F$ of functions which admit approximation by polynomials in some sense and try to define the value $f(A)$ for $f \in F$ as follows:

$$f(A) := \lim_{n\to\infty} P_n(A) \quad \text{where } P_n \to f \text{ when} \quad n \to \infty. \tag{11}$$

Here we have in mind that $P_n \to f$ in the sense of $F$ and $P_n(A) \to f(A)$ elementwise. This definition make sense if and only if the condition $P_n \to 0$ in the sense of $F$ implies $P(A) \to 0$ elementwise. Of course, it depends on the matrix $A$ and of the definition of a limit in $F$.

**Example 3** *Fix a finite set $S \subset \mathbb{R}$ and say that a sequence of functions $\{f_n\}$ is convergent if it converges at any point $x \in S$.*

*Then the definition* (11) *make sense for any matrix $A$ with a simple spectrum Spec $A \subset S$. Moreover, if $A = Q^{-1}DQ$, $D = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_k)$, then $f(A) = \operatorname{diag}\Big(f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_k)\Big)$.*

**Example 4** *Call a sequence of functions $\{f_n\}$ convergent if the sequences $\{f_n^{(i)}(x)$, $0 \le i \le n-1$, $x \in S$, are convergent.*

*Then the expression $f(A)$ make sense for any $(n-1)$-smooth function $f$ and for any matrix of size $\le n$ with a spectrum Spec $A \subset S$.*

**Example 5** *A function $f(x) = \frac{1}{1+x^2}$ is infinitely smooth and on any closed segment can be uniformly approximated by polynomials together with any number of derivatives.*
*However, for $A = \begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix}$ the value $f(A)$ is not defined since $1 + A^2 = 0$.*

## 1.5 Functions of Hermitian matrices

The examples in the end of the previous section show that the notion of a function of a matrix is rather delicate and handling with care is required. In the same time, there is a remarkable class of matrices $A$ for which the expression $f(A)$ is defined practically for all reasonable functions on $\mathbb{R}$. It is the class of *Hermitian matrices*, i.e. such complex matrices $A$ which satisfy

$$A = A^*, \quad \text{or} \quad a_{i,j} = \overline{a}_{j,i}. \tag{12}$$

In particular, all real symmetric matrices and pure imaginary antisymmetric matrices are Hermitian.

**Exercise 7** *Let a sequence* $\{P_n\}$ *of polynomials tends to 0 at any point* $x \in \mathbb{R}$*. Prove that* $P_n(A) \to 0$ *for any Hermitian matrix* $A$*.*

This statement remains true for matrices of infinite order corresponding to the so-called Hermitian operators in a Hilbert space. It plays a basic role in the mathematical model of quantum mechanics, because in this theory the physical observables are modelled by Hermitian operators.

We have, however, to make precise, what we understand under a matrix of infinite order and how to handle them. We are staying on a geometric base and consider matrices as operators acting on vectors. So, we must make clear, what is infinite vector. By the analogy with a finite-dimensional case, we could call an infinite real (or complex) vector any sequence $\curvearrowleft = (x_1, x_2, \ldots, x_n, \ldots)$ where $x_i$ are real (or complex) numbers. These vectors indeed form an infinite-dimensional vector space. But if we want to study geometric notions, such as length, angle, orthogonality, we have to restrict the class of sequences considered. Namely, call a sequence $\curvearrowleft = \{x_k\}_{k\geq 1}$ *admissible* if the series $\sum_{k\geq 1} |x_k|^2$ is convergent. The basic properties of admissible sequences we collect in the following

**Exercise 8** *Prove that*

*a) For any two admissible sequences* $\mathbf{x} = \{x_k\}_{k\geq 1}$ *and* $\mathbf{y} = \{y_k\}_{k\geq 1}$ *the series* $\sum_{k\geq 1} x_k \overline{y}_k$ *is convergent.*

*b) The admissible (real or complex) sequences form a (real or complex) vector space under the usual operations of addition and multiplication by a number.*

*c) Define the scalar (or dot-) product of admissible sequences and the length of* $\mathbf{x}$ *by the formulas*

$$(\mathbf{x}, \mathbf{y}) = \sum_{k\geq 1} x_k \overline{y}_k, \qquad |\mathbf{x}| = \Big(\sum_{k\geq 1} |x_k|^2\Big)^{\frac{1}{2}} \tag{13}$$

*Then the following Cauchy inequality holds:*

$$|(\mathbf{x}, \mathbf{y})| \leq |\mathbf{x}| \cdot |\mathbf{y}|. \tag{14}$$

The space of all admissible sequences is called a (real or complex) Hilbert space and is usually denoted by $H$ in honor of David Hilbert who introduced and used this space for studying the integral equations of mathematical physics. More details you can find in [19] and in [9] where you can find several beautiful problems from the geometry of Hilbert space.

Now we can tell which infinite matrices we shall consider: those, which define a continuous linear operators in a Hilbert space. Unfortunately, it is rather difficult to describe this class of matrices in terms of the matrix entries. A necessary, but not sufficient, condition is the existence of a constant $C$ such that $\sum_i |a_{i,j}|^2 \leq C$ for all $j$ and $\sum_j |a_{i,j}|^2 \leq C$ for all $i$. A sufficient, but not necessary, condition is the convergence of the double series $\sum_{i,j} |a_{i,j}|^2$.

What can be said about the spectrum of an infinite Hermitian matrix?[5]

It is easy to give an example of a (diagonal) matrix which has a countable set of eigenvalues. More instructive is the following example, which shows that the spectrum can contain a whole segment of the real axis.

**Example 6** *Consider an infinite matrix $A$ with entries*

$$a_{i,j} = \begin{cases} 1 & \textit{if} \quad |i-j| = 1 \\ 0 & \textit{otherwise.} \end{cases}$$

*To understand the properties of the operator defined by this matrix, we establish the correspondence between admissible sequences $\mathbf{x} = \{x_k\}_{k\geq 1}$ and $2\pi$-periodic odd functions on a real line with a coordinate $t$, by the formula*

$$\{x_k\}_{k\geq 1} \longleftrightarrow \sum_{k=1}^{\infty} x_k \sin kt. \tag{15}$$

*It can be shown that the image $\widetilde{H}$ of the space $H$ under this correspondence consists of all $2\pi$-periodic odd real functions $\varphi$ on $\mathbb{R}$ which satisfy the condition*

$$\int_0^\pi |\varphi(t)|^2 dt < \infty \quad (\textit{in the sense of Lebesgue integral}).$$

*The scalar product in $H$ goes to a scalar product in $\widetilde{H}$ given by*

$$(\varphi,\, \psi) = \frac{2}{\pi} \int_0^\pi \varphi(t)\psi(t) dt.$$

[5]The reader brought up on the rigorous principles of mathematical analysis will, probably, refuse to discuss this question in the lack of an accurate definition of the spectrum. I think, however that it is more important to have an intuitive idea of spectrum than to learn by rote its definition.

**Exercise 9** *Let* $\varphi(t) = t$ *for* $-\pi \leq t < \pi$ *and be extended further by* $2\pi$*-periodicity. Show that* $\varphi \in \widetilde{H}$*, find its preimage* $\mathbf{x} \in H$*. What means in this case the equality* $|\varphi|^2 = |\mathbf{x}|^2$*?*

**Exercise 10** *Show that the linear operator in* $H$ *given by the matrix* $A$ *goes under correspondence* (15) *to the operator of multiplication by the function* $a(t) = 2\cos t$ *in the space* $\widetilde{H}$*.*

Hint: Use the equality $2cos\, t \sin nt = \sin (n-1)t + \sin (n+1)t$.

It follows that the expression $f(A)$ make sense for any functions on the segment $[-2, 2]$ and that corresponding operator in $\widetilde{H}$ is the multiplication by the function $f\big(a(t)\big)$.

**Exercise 11** *Show that the matrix* $A$ *has no eigenvectors in* $H$*.*

We see that in the infinite-dimensional case the definition of the spectrum as the collection of eigenvalues is no good at all. The true definition, working in both, finite-dimensional and infinite-dimensional cases, and ensuring the validity of the analogues of theorems (2) and (3), is given in textbooks on functional analysis. Namely, a point $\lambda$ *does not* belong to Spec $A$ if the operator $A - \lambda \cdot 1$ has a continuous inverse operator $R_\lambda(A) := (A - \lambda \cdot 1)^{-1}$, the so-called *resolvent* for $A$.

## 2 Continuous matrices and von Neumann factors

What is rational is real; And what is real is rational.

Hegel, Preface to "Elements of the Philosophy of Right"

### 2.1 Infinite matrices

The notion of matrix admits several variants of transition to infinity. One of them we examined in Section 1. There the matrix elements were numerated by indices $i$, $j$ which run through the set $\mathbb{N}$ of natural numbers.[6]

Another variant is to consider index as a continuous variable taking, for instance, real values. Then vectors become functions $\varphi$ of a real variable $t$, matrices become functions $A$ of two real variables $t$, $s$, summation over a discrete index becomes an integration and linear operators become integral operators of the form

$$A\varphi(t) = \int_{\mathbb{R}} A(t, s)\varphi(s)ds \tag{16}$$

To make things precise, we have to tell, which kind of functions we consider and what kind of integration we use. It turns out that after the most natural and convenient choice of all specifications, we come to the same theory of Hilbert spaces which we discussed above. The point is that one can establish a correspondence between a space of functions and a space of sequences so that all operations in Hilbert space: addition, multiplication by a number, scalar product and limit are preserved. Such a correspondence we constructed in Example (6).

In spite of its simplicity, the correspondence in question is a very strong and deep result with many application in mathematics and mathematical physics.

There is one more variant to construct infinite matrices. It was used by John von Neumann about 70 years ago to construct remarkable infinite-dimensional algebras, so-called *von Neumann factors*. This theory leads to a surprising generalization of vector spaces where the dimension can take

[6] Of course, we could use instead of $\mathbb{N}$ any other countable set. Often it is convenient to use the set $\mathbb{Z}_+$ of all non-negative integers, or the set $\mathbb{Z}$ of all integers, or $n$-dimensional lattice $\mathbb{Z}^n$ and the like.

any real values. At that time this discovery was not further developed and was considered rather as curious example like a non-measurable set. Now the theory of von Neumann factors is one of central domain in functional analysis and in quantum field theory, see e.g. [2].

I explain here several facts from this theory in some details, because I do not know any popular exposition of them.

**Example 7** *Consider matrices of a growing size $n$. We shall write them as $n \times n$ tables of the fixed size but with smaller and smaller boxes. It is convenient to put $n = 2^k$, so that every algebra $\mathcal{M}_k = \mathrm{Mat}_{2^k}$ can be embedded in the next algebra $\mathcal{M}_{k+1}$. Let us think, what we get in the limit $k \to \infty$.*

Consider the unit square on the plane and draw the main diagonal from upper left to the lower right corner and also the parallel lines which divide each side of the square into $2^k$ equal parts. Denote the set obtained by $X_k$. Then divide all lines in $X_k$ on equal parts with horizontal and vertical projections of the size $2^{-k}$. Then the whole set $X_k$ will contain $2^{2k}$ parts. Consider the set $V_k$ of complex-valued functions on $X_k$ which are constant on each of its part; we call such functions *$k$-locally constant.* It is a $2^{2k}$-dimensional complex vector space which is naturally isomorphic to the space $\mathcal{M}_k = \mathrm{Mat}_{2^k}(\mathbb{C})$. Finally, consider the set $X := \bigcup_{k\geq 1} X_k$ and identify $V_k$ with the space of functions on $X$ which are $k$-locally constant and vanish outside $X_k$. Then $V_k \subset V_{k+1}$ and the space $V = \bigcup_{k\geq 1} V_k$ has a structure of infinite-dimensional algebra

$$\mathcal{M} := \bigcup_{k\geq 1} \mathcal{M}_k. \tag{17}$$

One can ask, how we define functions from $\mathcal{M}$ in the boundary points of the form $(r_1, r_2)$ where $r_i$ are dyadic rational numbers $\frac{k_i}{2^{n_i}}$. There are two reasonable answers:

1. It does not matter for the further constructions because there are only countable set of these boundary points; and we shall consider only equivalence classes of functions: $f_1 \equiv f_2$ if they coincide almost everywhere.

2. Split every dyadic rational points into two points. Recall that in dyadic system each dyadic rational number can be written in two different ways, e.g. 0.011111... and 0.100000.... So, if we will consider any infinite 2-adic fraction as individual number, than we get the desired splitting. This operation turns the segment [0, 1] into totally disconnected set, known as *Cantor set.* The inverse operation of gluing points defines the so-called

*Cantor ladder* (or *Devil ladder*), a function which maps the Cantor set *onto* the segment [0, 1].

The set $X$ after this modification also becomes a totally disconnected set $\widetilde{X}$ and all functions from $\mathcal{M}$ become continuous. Moreover, the set $\widetilde{X}$ is a direct product of two topological groups.

Namely, let $\Gamma$ be the group of all infinite sequences with elements from $\{0, 1\}$ with the group law, given by componentwise addition $\mod 2$. It is just a complete direct product of a countable set of copies of $\mathbb{Z}_2$. as a topological space, $\Gamma$ is homeomorphic to Cantor set.

On the other hand there exists a direct sum (or restricted product) $\Gamma_0$ of a countable set of copies of $\mathbb{Z}_2$. It is a countable discrete group, consisting of those infinite sequences with elements from $\{0, 1\}$, which contain only finite number of 1's.

The set $\widetilde{X}$ is identified with $\Gamma_0 \times \Gamma$: the elements of the first group counts the "diagonals" and elements of the second group label the points on a given diagonal. Here we consider as a diagonal the subset of points $x$, $y$ from the unit square given by equation

$$x - y \;=\; const \mod 1.$$

The construction below make sense for more general pairs of groups.

Come back to the space $\mathcal{M}$. We can introduce in it two different norms. First, we can define a scalar product in $\mathcal{M}$ putting

$$(\varphi, \psi) = \int_X \varphi(x)\overline{\psi(x)}dx \tag{18}$$

where $dx$ denote the measure on $X$ which on every segment of a diagonal is equal to the length of the horizontal projection.

**Exercise 12** *Show that on the subspace $\mathcal{M}_k \subset \mathcal{M}$ the scalar product* (18) *can be given by the formula*

$$(A, B) = 2^{-k}\mathrm{tr}\,(AB^*) = \frac{\mathrm{tr}\,(AB^*)}{\mathrm{tr}\,(1)} \tag{19}$$

*where* tr *denotes the trace of a matrix and* $^*$ *means Hermitian conjugation.*

As usual, when you have a scalar product, you can define the length of a vector by $|\varphi| = \sqrt{(\varphi, \varphi)}$ the distance $d(\varphi_1, \varphi_2) = |\varphi_1 - \varphi_2|$ and the limit: $\varphi_n \to \varphi \iff d(\varphi_n, \varphi) \to 0$. The completion of $\mathcal{M}$ with respect to this distance we call $H$; it is a complex Hilbert space. In our realization $H$ coincides with the space $L^2(X, dx)$ of square-integrable (in Lebesgue sense) functions on $X$. The inner product is given by the same formula (18) where integral is understood in Lebesgue sense.

## 2.2 Construction of an algebra $\mathcal{C}$

Recall now that $\mathcal{M}$ is an associative algebra. Indeed, any two elements $\varphi_1$, $\varphi_2$ from $\mathcal{M}$ belong to some $\mathcal{M}_k$. Thus, they are identified with matrices from $\mathrm{Mat}_{2^k}(\mathbb{C})$ and can be multiplied as such. This multiplication law can be given by the formula

$$(\varphi_1 \cdot \varphi_2)(x,\, y) = \sum_t \varphi_1(x,\, t)\varphi_2(t,\, y) \tag{20}$$

where summation is over a countable set of such $t \in [0,\, 1]$, for which $x - t$ and $t - y$ are dyadic fractions (practically, for $\varphi_i \in \mathcal{M}$ the summation is over the finite set of $t$ such that $x - t$ and $t - y$ have denominator $2^k$). Note that (20) is similar to the ordinary rule of matrix multiplication.

Now we introduce in $\mathcal{M}$ another norm putting

$$\|\varphi\| = \sup_{\psi \neq 0} \frac{|\varphi \cdot \psi|}{|\psi|} \tag{21}$$

where $|\cdot|$ is the length defined above. Note, that this definition implies (if we put $\psi = 1$) that $\|\varphi\| \geq |\varphi\|$.

**Exercise 13** *Show that for matrices $A$ of second order (or, for elements $\mathcal{M}_2$)*

$$|A|^2 = \mathrm{tr}\,(AA^*), \qquad \|A\|^2 = \frac{|A|^2 + \sqrt{|A|^4 - 4|\det A|^2}}{2}.$$

**Exercise 14** *Show that for matrices of format $n \times n$*

$$|A| \;\leq\; \|A\| \;\leq\; |A| \cdot \sqrt{n}. \tag{22}$$

The completion of $\mathcal{M}$ with respect to the norm $\|\cdot\|$ we denote by $\mathcal{C}$. From this definition follows that $\mathcal{C}$ can be considered as an algebra of continuous operators in $H$ which is *symmetric* in the sense that if $A \in \mathcal{C}$, then $A^* \in \mathcal{C}$. Moreover, $\mathcal{C}$ is closed with respect to the norm $\|\cdot\|$, which coinsides with the ordinary operator norm.

Since the convergence in $\mathcal{C}$ is stronger than in $H$, we can consider $\mathcal{C}$ as a subspace in $H$ (dense in the topology of $H$). In our model $\mathcal{C}$ consists of certain functions on $\widetilde{X}$ which are continuous and going to zero at infinity. (Indeed, functions from $\mathcal{M}$ have these properties and convergence in norm implies uniform convergence which preserve them.) The precise description of this class of functions would be interesting but I do not know the answer.

## 2.3 Construction of a factor

Finally, we introduce in $\mathcal{M}$ one more definition of convergence which is called *strong operator convergence*. Namely, we say that a sequence $A_n$ of continuous operators in a Hilbert space $H$ strongly converges to an operator $A$ if for any vector $v \in H$ we have $|A_n v - Av| \to 0$.

This kind of convergence is weaker than norm convergence, but stronger than convergence in $H$. The completion of $\mathcal{M}$ with respect to this convergence we denote by $\mathcal{W}$. Unfortunately, we do not know yet the description of $\mathcal{W}$, interpreted as a function space on $\tilde{X}$. It is certainly is between the space $L_2(\widetilde{X}, dx)$ of square integrable functions and the space $C_0(\widetilde{X})$ of continuous functions, vanishing at infinity.

**Lemma 2** *The space $\mathcal{W}$ consists of all elements $\varphi \in H$ for which*

$$\|\varphi\| := \sup_{\psi \in \mathcal{C} \setminus \{0\}} \frac{|\varphi \cdot \psi|}{|\psi|} < \infty. \tag{23}$$

**Corollary 1** *For $\varphi \in \mathcal{W}$ and $\psi \in H$ the products $\varphi \cdot \psi$ and $\psi \cdot \varphi$ are defined in $H$. Moreover,*

$$|\varphi \cdot \psi| \leq \|\varphi\| \cdot |\psi| \quad \textit{and} \quad |\psi \cdot \varphi| \leq |\psi| \cdot \|\varphi\|. \tag{24}$$

So, the algebra $\mathcal{W}$ has two realizations as an algebra of operators in $H$: the algebra $L(\mathcal{W})$ of all operators of left multiplication $\psi \mapsto \varphi \cdot \psi$ and the algebra $R(\mathcal{W})$ of operators of right multiplication: $\psi \mapsto \psi \cdot \varphi$. Here $\varphi \in \mathcal{W}$, $\psi \in H$. Moreover, these realizations are isometric: the norm of element $\varphi \in \mathcal{W}$ is equal to norms of operators $L(\varphi)$ and $R(\varphi)$.

**Corollary 2** *For $\varphi \in \mathcal{W}$ and $\psi \in H$ assume that*

$$|\varphi \cdot \psi| \leq |\varphi| \cdot a \quad \textit{and} \quad |\psi \cdot \varphi| \leq b \cdot |\varphi|. \tag{25}$$

*Then $\psi \in \mathcal{W}$ and $\|\psi\| \leq \min(a, b)$.*

## 2.4 von Neumann algebras

For any non-empty set $S$ of operators in a Hilbert space $H$ we define by $S^!$ the *commutant* of $S$ which consists of all operators, commuting with every operator from $S$. von Neumann proved the following remarkable

**Theorem 4 (von Neumann)** *If the set $S$ is symmetric (i.e. together with any operator $A$ contains the adjoint operator $A^*$), then $S^{!!!} = S^!$ and $S^{!!}$ is the closure of the algebra, generated by $S$ in the sense of strong operator topology.*

**Exercise 15** *Prove this theorem for finite-dimensional Hilbert spaces.*

**Definition 1** *An von Neumann algebra is an operator algebra satisfying* $A^{!!} = A$.

Come back to our example. The first important property of algebras $L(\mathcal{W})$ and $R(\mathcal{W})$ is described by the lemma.

**Lemma 3**

$$L(\mathcal{W})^! = R(\mathcal{W}), \qquad R(\mathcal{W})! = L(\mathcal{W}).$$

*Proof.* *Let* $C$ *be an operator in* $H$ *and let* $C \in L(\mathcal{W})^!$. *Then for any* $B \in \mathcal{M} \subset H$ *we have*

$$C(B) = C(B \cdot 1) = CL(B)(1) = L(B)C(1) = BC(1) = R\big(C(1)\big)(B),$$

*hence,* $C = R\big(C(1)\big)$ *and we are done, if only we knew that* $C(1) \in \mathcal{W}$. *But this can be deduced from corollary (2).*

Thus, the algebras $L(\mathcal{W})$ and $R(\mathcal{W})$ are mutual commutants and are von Neumann algebras.

The second basic property is that the intersection $L(\mathcal{W}) \bigcap R(\mathcal{W})$ consists only of scalar operators. We leave to the reader to prove this property.

On the other hand, for any von Neumann algebra $\mathcal{A}$ it is easy to see, that $\mathcal{A} \bigcap \mathcal{A}^!$ coincides with the common center of $\mathcal{A}$ and $\mathcal{A}^!$. The von Neumann algebras with trivial center (i.e. consisting of scalar operators only) von Neumann himself called *factors*. This name comes from the following remarkable fact: any von Neumann algebra can be realized as a continuous product of factors. So, factors play the role of prime numbers or simple groups, or irreducible representations.

**Exercise 16** *Let* $H$ *be a finite-dimensional Hilbert space and* $\mathcal{W}$ *is a symmetric algebra of operators in* $H$*, containing 1.*

*a) Prove that* $\mathcal{W}^{!!} = \mathcal{W}$*, i.e.* $\mathcal{W}$ *is a von Neumann algebra.*

*b) Assume that* $\mathcal{W}$ *is a factor, i.e. the center of* $\mathcal{W}$ *consists of scalars. Show that* $H$ *can be realized as a space of matrices of the format* $m \times n$ *so that* $\mathcal{W} \simeq \mathrm{Mat}_m(\mathbb{C})$, $\mathcal{W}^! \simeq \mathrm{Mat}_n(\mathbb{C})$*, the first algebra acts on* $H$ *by left multiplication and the second by right multiplication.*

The factors $\mathcal{W}$ and $\mathcal{W}!$ from the exercise are called *factors of type* $I_m$ and $I_n$ respectively.

This result admits a partial generalization to the infinite-dimensional case. Namely, if $\mathcal{W}$ is a factor in a Hilbert space $H$ and if $\mathcal{W}$, as a topological algebra is isomorphic to the algebra $L(H_1)$ of all continuous linear operators in a Hilbert space $H_1$, then the algebra $\mathcal{W}^!$ is isomorphic to $L(H_2)$ for some Hilbert space $H_2$ and $H$ can be identified with a Hilbert tensor product $H_1 \boxtimes H_2$. The factor $\mathcal{W}$ is called a *factor of type* $I_\infty$. The dual factor $\mathcal{W}!$ belongs to the type $I_n$ or $I_\infty$ depending on the dimension of $H_2$.

The exercise above shows that only factors of type I occur in finite-dimensional case.

## 2.5 Relative dimensions

The von Neumann discovery is that there are another types of factors which are non-isomorphic to $L(H)$. We show below that ours $L(\mathcal{W})$ and $R(\mathcal{W})$ are examples of such factors of new type. For this goal we introduce the notion of *relative dimension* of two subspaces in $H$, invented by von Neumann.

Let $\mathcal{W}$ be a factor acting in $H$. Consider the set of all closed subspaces in $H$ which are stable with respect to $\mathcal{W}$. We call such spaces *admissible.*

**Exercise 17** *Let $H_1$ be a closed subspace in $H$ and $P$ be the orthoprojector from $H$ to $H_1$. Show that $H_1$ is admissible iff $P \in \mathcal{W}^!$.*

Below we often identify the subspaces with corresponding orthoprojectors.

Let us call two admissible subspaces $H_1$, $H_2$ *equivalent* if there exists an element $u \in \mathcal{W}^!$ such that

$$u^*u = P_1, \quad uu^* = P_2 \quad \text{where } P_i \text{ is the orthoprojector to } H_i. \tag{26}$$

In this case we write $H_1 \approx H_2$.

**Exercise 18** *Show that equations* (26) *are equivalent to the statement that $u$ maps $H_1$ isometrically onto $H_2$ and annihilates $H_1^\perp$ while the adjoint operator $u^*$ maps $H_2$ isometrically onto $H_1$ and annihilates $H_2^\perp$.*

We shall say that an admissible subspace $H_1$ is *bigger* than another admissible subspace $H_2$ if $H_1$ has a part, equivalent to $H_2$. An admissible subspace $H_1$ is called *infinite* if it contains a proper subspace, equivalent to the whole $H_1$; otherwise, it is called *finite.*

The most interesting are those factors for which the whole space $H$ is finite (being, of course, infinite-dimensional in the ordinary sense). Such factors are called *factors of type* $II_1$. We consider in more details the geometry related to these factors.

**Theorem 5 (von Neumann)** *Let $\mathcal{W}$ is a factor acting in a Hilbert space $H$. For a pair of admissible subspaces $H_1, H_2 \subset H$ exactly one of the following is true:*

*1. $H_1$ is bigger than $H_2$, but $H_2$ is not bigger than $H_1$;*
*2. $H_2$ is bigger than $H_1$, but $H_1$ is not bigger than $H_2$;*
*3. $H_1$ and $H_2$ are equivalent.*

The sketch of the proof. There are two alternatives:
$A_1$) $H_1$ contains a part equivalent to $H_2$;
$A_2$) $H_1$ does not contains a part equivalent to $H_2$.
and
$B_1$) $H_2$ contains a part equivalent to $H_1$;
$B_2$) $H_2$ does not contains a part equivalent to $H_1$.
So, we have together 4 possibilities: $A_1\&B_1,\ A_1\&B_2,\ A_2\&B_1,\ A_2\&B_2$.

**Lemma 4** *In the case $A_1\&B_1$ we have $H_1 \approx H_2$.*

The proof is analogous to the proof of the Cantor-Bernstein theorem in the set theory. Namely, let $u \in \mathcal{W}^!$ be an operator which maps isometrically $H_1$ to the part of $H_2$ and vanishes on $H_1^\perp$. Also, let $v \in \mathcal{W}^!$ be an operator which maps isometrically $H_2$ to the part of $H_1$ and vanishes on $H_2^\perp$. Then we have two decreasing chains of admissible subspaces:

$$\begin{array}{l} H_1 \supset vH_2 \supset vuH_1 \supset vuvH_2 \supset vuvuH_1 \supset uvuvuH_2 \supset \dots \\ H_2 \supset uH_1 \supset uvH_2 \supset uvuH_1 \supset uvuvH_2 \supset uvuvuH_1 \supset \dots \end{array} \tag{27}$$

Denote by $V_1$ the subspace $H_1 \ominus vH_2 := H_1 \bigcap (vH_2)^\perp$ and by $V_2$ the subspace $H_2 \ominus uH_1$. Then we obtain

$$\begin{array}{l} H_1 = V_1 \oplus vV_2 \oplus vuV_1 \oplus vuvV_2 \oplus \dots \oplus V_1^\infty \\ H_2 = V_2 \oplus uV_1 \oplus uvV_2 \oplus uvuV_1 \oplus \dots \oplus V_2^\infty \end{array} \tag{28}$$

where

$$V_1^\infty := \bigcap_{k \geq 0} (vu)^k H_1, \quad V_2^\infty := \bigcap_{k \geq 0} (uv)^k H_2.$$

Look on the equivalence classes of all these subspaces. It is clear that in the first sequence in (28) the subspaces are equivalent alternatively to $V_1$ and $V_2$, while in the second sequence they are equivalent alternatively to $V_2$ and $V_1$. Also, $uV_1^\infty = V_2^\infty$ and $vV_2^\infty = V_1^\infty$ and $uv|_{V_2^\infty} = Id$, so that $V_1^\infty$ and $vV_2^\infty$ are equivalent.

It is now easy to establish equivalence between $H_1$ and $H_2$

**Lemma 5** *The case $A_2\&B_2$ never happens.*

For any unitary operator $u \in \mathcal{W}^!$ consider the subspaces $V_1 = H_1 \bigcap uH_2$ and $V_2 = H_2 \bigcap u^{-1}H_2$. It is clear that $V_i \subset H_i$ and $V_1 \approx V_2$. Let us show that these subspaces are non-zero for some $u \in \mathcal{W}^!$. Assume the contrary. Then the space $\mathcal{W}^!\, H_2$ has zero intersection with $H_1$, hence, is a non-trivial subspace invariant with respect to $\mathcal{W}^!$. But this subspace is also admissible, hence, invariant with respect to $\mathcal{W}$. Therefore, the projector to this subspace belong to $\mathcal{W} \bigcap \mathcal{W}^! = \{\mathbb{C} \cdot 1\}$, a contradiction.

So, any two admissible subspaces $H_1$, $H_2$ have non-zero equivalent parts. Using Zorn's Lemma, one can show that there is a maximal element in the set of all pairs of equivalent parts. From the lemma, this maximal element must contain either $H_1$ or $H_2$.

So, any two admissible subspaces are comparable: either they are equivalent, or one of them is bigger that another. Actually these possibilities are not exclude one another. But then a subspace is equivalent to its proper part, hence is infinite. The factor is called *of type II* if any admissible subspace $H_1$ contains a finite subspace $V \subset H_1$.

The most interesting are those factors of type II for which the whole space $H$ is finite (being, of course, infinite-dimensional in the ordinary sense). Such factors are called *factors of type $II_1$*. We consider in more details the geometry related to these factors.

First of all, the equivalence classes of admissible subspaces are ordered: $[H_1] > [H_2]$ if $H_1$ contains a part, equivalent to $H_2$. (Here $[H_1]$ is the equivalence class of $H_1$.) Moreover, we can compare the relative size of two admissible subspaces exactly how in Euclidean geometry we compare the relative size of two segments. Namely, let $[H_1] > [H_2]$. Denote by $n_1$ the maximal number $n$ such that $H_1$ contains orthogonal sum of $n$ subspaces $V_i$, $1 \leq i \leq n$, each of which is equivalent to $H_2$.

**Exercise 19** *Show that* $1 \leq n_1 < \infty$.

Then denote by $H_3$ the orthogonal complement in $H_1$ to the $\oplus_{i=1}^{n_1} V_i$.

**Exercise 20** *Show that* $[H_3] < [H_2]$.

Denote by $n_2$ the maximal number $n$ such that $H_2$ contains orthogonal sum of $n$ subspaces $W_i$, $1 \leq i \leq n$, each of which is equivalent to $H_3$. And so on...

It can happen that after several steps we come to the case $H_{k+1} = \{0\}$. Then both $H_1$ and $H_2$ can be presented as orthogonal sums of parts, equivalent to $H_k$, so that $H_i$ contains $M_i$ such parts.

**Exercise 21** *Show that*

$$r(H_1,\,H_2) := \frac{M_1}{M_2} = n_1 + \cfrac{1}{n_2 + \cfrac{1}{n_2 + \cfrac{1}{\ddots + \cfrac{1}{n_k}}}}$$

We call $r(H_1,\,H_2)$ the *relative dimension* of $H_1$ and $H_2$.

It is possible, though, that the process never ends. Then $H_1$ and $H_2$ are *non-commensurable* and their relative dimension is an irrational number $r(H_1,\,H_2)$, given by an infinite continuous fraction.

## 2.6 Relative trace

.

The notion of relative dimension allows to define the so-called *relative trace* on the algebra $\mathcal{W}^!$. Namely, for any orthoprojector $P$ in $\mathcal{W}^!$ we define

$$\operatorname{tr} P = r(PH,\,H) \tag{29}$$

and extend this definition to the whole algebra by linearity and continuity in strong operator topology.

We now show that the algebra $L(\mathcal{W})$ constructed in Section 2.3, is a factor of type $II_1$. Indeed, we know that $L(\mathcal{W})^! = R(\mathcal{W})$

**Exercise 22** *Show that for an operator in* $L(\mathcal{W})^! = R(\mathcal{W})$ *corresponding to a function* $\varphi$ *on* $X$ *the relative trace can be computed by the formula*

$$\operatorname{tr} \varphi = \int_0^1 \varphi(t,\,t)dt. \tag{30}$$

**Exercise 23** *Find in the space* $H$

*a) two orthogonal and equivalent admissible subspaces;*

*b) three pairwise orthogonal and equivalent admissible subspaces.*

*c) establish explicitly the equivalence of subspaces in question.*

**Example 8** *Quantum torus* $\mathbb{T}_q$.

*The title of this example reflects the general tendency to relate to quantum theory every transition from a commutative objects to non-commutative ones.*

It is well-known that any compact topological space is determined by the algebra $C(X)$ of continuous functions on $X$ and any smooth compact manifold $M$ is determined by the algebra $C^\infty(M)$ of smooth functions on $M$.[7] The corresponding quantum objects arise if we replace the commutative algebra $A = C^\infty(M)$ by some non-commutative algebra $A_q$

In the case $M = \mathbb{T}^2$ the algebra $A$ consists of functions

$$f(x, y) = \sum_{k,l} c_{k,l} e^{2\pi i(kx+ly)}, \tag{31}$$

where summation is over 2-dimensional lattice $\mathbb{Z}^2$, and the coefficients satisfy for every $N \in \mathbb{N}$ the conditions:

$$|c_{k,l}| \le C_N \cdot (1 + k^2 + l^2)^{-N}. \tag{32}$$

Let now $\widetilde{A}_q$ be an associative algebra with involution, generated by the unit and by two elements $u$, $v$ satisfying the relations

$$u^*u = uu^* = 1, \qquad v^*v = vv^* = 1, \qquad uv = qvu, \tag{33}$$

where $q = e^{2\pi i\tau}$ is a complex number with $|q| = 1$. By $A_q$ we denote the completion of $\widetilde{A}_q$, which consists of formal sums

$$\sum_{k,l} c_{k,l}\, u^k v^l, \tag{34}$$

where the coefficients satisfy the conditions (32). We consider $A_q$ as an algebra of smooth functions on the *quantum torus* $\mathbb{T}^2_q$. It is clear that for $q = 1$ the quantum torus becomes the ordinary one.

From the algebra $A_q$ one can construct a Hilbert space $H$, the algebra $W$ and two its representations $L$ and $R$ in the space $H$ along the lines used above, so that

$$L(W) = R(W)^!, \qquad R(W) = L(W)^!, \qquad L(W) \bigcap R(W) = \mathbb{C} \cdot 1.$$

It turns out, that the dual factors $L(W)$ and $R(W)$ belong to type $I_\infty$ when $\tau \in \mathbb{Q}$ (i.e. when $q$ is a root of unity) and to type $II_1$ otherwise.

**Exercise 24** *Show that $II_1$-factors of example 8 are isomorphic to factors constructed before.*

[7] For instance, the points of $M$ can be reconstructed as maximal ideals in $C^\infty(M)$, smooth vector fields on $M$ are just derivations of $C^\infty(M)$ and so on (see $ 7).

**Example 9** *Quantum spheres $S_q^2$ and $S_q^3$.*

*It is well-known that the sphere $S^3$ has a group structure (see \$ 4) and the sphere $S^2$ can be considered as homogeneous $S^3$-space. All these statements can be formulated in terms of smooth functions on $S^2$ and $S^3$, never mentioning the points. Then we can build the quantum analogues by replacing $C^\infty(S^2)$ and $C^\infty(S^3)$ by appropriate non-commutative algebras*

For example, the fact that $S^3$ acts on $S^2$ is described by a smooth map

$$S^3 \times S^2 \longrightarrow S^2.$$

And this map is determined by an algebra homomorphism

$$C^\infty(S^2) \longrightarrow C^\infty(S^3 \times S^2) = C^\infty(S^2) \widehat{\otimes} C^\infty(S^2).$$

It turns out that there exist a non-commutative deformations $A_q$ of $C^\infty(S^3)$ and $B_q$ of $C^\infty(S^2)$ such that all necessary homomorphisms make sense. This allow us to speak about quantum group $S_q^3$ and of its homogeneous space $S_q^2$. Detailed description of these deformation was first done in [16].

Hints and answers.

# 3 What is supersymmetry?

> I realized that China and Spain is absolutely the same land, and only by ignorance are considered as different countries. I advise to everybody to write down on a paper Spain and you get China.
>
> N. V. Gogol, "Notes of a madman".

## 3.1 Symmetry and supersymmetry

In mathematics, as in all other sciences, the symmetry plays a very important role. Usually, the properties of symmetry are expressed using the notions of groups, homogeneous spaces and linear representations (see [8]). But in last quarter of century a new type of symmetry acquires bigger and bigger value, especially in mathematical physics and related domains of mathematics. It got a nickname SUSY from the physical term *supersymmetry.*

In mathematics, the supersymmetry means, in a sense an equality of rights for plus and minus, for odd and even, for symmetric and antisymmetric etc. In physics, this is a parallel between bosonic and fermionic theories.

In a few words, the ideology of supersymmetry can be formulated as follows. To every *ordinary* or *even* notion (definition, theorem, construction etc) there must correspond an *odd* analogue, which together with the original object form a *superobject*. And this superobject manifest in a higher level the features of the initial ordinary object.

In particular, the SUSY formalism requires the new kind of numbers. They differ from the ordinary ones by the property of multiplication: it is not commutative, but anticommutative: $xy = -yx$. In particular, $x^2 = 0$. These new numbers must be used as widely and with same rights as ordinary numbers. E.g., they can play the role of coordinates of a vector, arguments of a function, local coordinates on a manifold (see, e.g.,[14]).

For some notions the odd analogues are rather evident; for others they are more complicated and sometimes unexpected. I bring here only a few elementary examples, leaving to the active reader to extend this list. Try, for example, to answer the question: what means the SUSY for a problem you are working now (or have worked recently)?

## 3.2 Supersymmetry in algebra

Let us start with the linear algebra. The superanalogue of a vector space is a $\mathbb{Z}_2$-graded vector space $V$ decomposed into direct sum: $V = V_{\bar{0}} \oplus V_{\bar{1}}$. Here $\bar{0}$, $\bar{1}$ are elements of $\mathbb{Z}_2 := \mathbb{Z}/2\mathbb{Z}$; vectors from $V_{\bar{0}}$ are called even, vectors from $V_{\bar{1}}$ are called odd. The pair ($\dim V_{\bar{0}}$, $\dim V_{\bar{1}}$) is called a *dimension* of $V$.[8]

**Exercise 25** *Define a direct sum and a direct product of $\mathbb{Z}_2$-graded vector spaces so that their dimensions were correspondingly added and multiplied. (You have also define the addition and multiplication rules for dimensions.)*

**Exercise 26** *Define the index of a $\mathbb{Z}_2$-graded vector space $V = V_{\bar{0}} \oplus V_{\bar{1}}$ the number $i(V) := \dim V_{\bar{0}} - \dim V_{\bar{1}}$. Show that*

$$i(V \oplus W) = i(V) + i(W), \qquad i(V \otimes W) = i(V) \cdot i(W).$$

**Remark 1** *Sometimes the index itself is considered as a right superanalogue of dimension. The only disadvantage of index in the role of superdimension is that two spaces with equal indices are not necessarily isomorphic.*

[8]More precisely, it should be called *superdimension* but I do not want to abuse the prefix "super".

**Remark 2** *The notion of the index of a $\mathbb{Z}_2$-graded vector space is long ago in use in algebraic topology. Indeed, the Euler characteristic, the Lefschetz number and other alternated sums in topology are all examples of indices.*

Pass now to linear operators. For two $\mathbb{Z}_2$-graded vector spaces $V$, $W$ we denote by $L(V, W)$ the set of all linear maps from $V$ to $W$ (disregarding the grading). This set has also the natural structure of a $\mathbb{Z}_2$-graded vector space. Namely write an operator $A \in L(V, W)$ as a block matrix

$$A = \begin{Vmatrix} A_{\bar{0}\bar{0}} & A_{\bar{0}\bar{1}} \\ A_{\bar{1}\bar{0}} & A_{\bar{1}\bar{1}} \end{Vmatrix} \tag{35}$$

where $A_{ij}$ is an operator from $V_i$ to $W_j$. The blocs along the main diagonal are even, the off-diagonal blocks are odd. So, if we denote by $p(x)$ the *parity* of an object $x$, taking values in $\mathbb{Z}_2$, then for an operator $A$ and a vector $v$ we have

$$p(Av) = p(A) + p(v). \tag{36}$$

This principle holds for all other multiplicative operations: multiplication of operators, tensor product of vector spaces and operators, inner product etc. So, we have

**Rule 1.** Under multiplication parities add.

Moreover, in all definitions, identities, equations etc which contain multiplicative operations, the additional factors $\pm 1$ arise according to[9]

**Rule 2.** If in an ordinary algebraic formula there are monomials with interchanged factors, then in corresponding superalgebraic formula each permutation of terms $x$ and $y$ must be accompanied by the extra factor $(-1)^{p(x)p(y)}$.

**Exercise 27** *Formulate the superanalogues of*

*a) commutativity:* $xy = yx$*;*

*b) associativity:* $(xy)z = x(yz)$*;*

*c) Jacobi identity:* $[x, [y, z]] + [y, z], x] + [z, x], y] = 0$.

An important notion in linear algebra is the notion of a trace $\operatorname{tr} A$ of an operator $A$. The characteristic property of the trace, which determines it up to scalar factor, is

$$\operatorname{tr} AB = \operatorname{tr} BA, \quad \text{or} \quad \operatorname{tr} [A, B] = 0.$$

[9] This formulation is taken from [14]. Physicists prefer more rich sentence: when something of parity $a$ flows past something of parity $b$, the sign $(-1)^{ab}$ arises.

In a superalgebra the commutator of homogeneous elements is given by the formula

$$[A,\, B] = AB - (-1)^{p(A)p(B)}BA. \qquad (37)$$

We define a *supertrace* of an operator $A$ written in the form (35) by the formula

$$\operatorname{str} A = \operatorname{tr} A_{00} - \operatorname{tr} A_{11}. \qquad (38)$$

**Exercise 28** *Show that* $str[A,\, B] = 0$ *for all* $A,\, B$.

**Exercise 29** *Check that the* $\mathbb{Z}_2$*-graded vector space* $L(V,\, V)$ *with the operation of supercommutator is a Lie superalgebra (i.e. the superanalogue of the Jacobi identity holds).*

The result of this and the previous exercises can be formulated as follows: the supertrace is a homomorphism of the Lie superalgebra $L(V,\, V)$ into $\mathbb{R}$ (with zero commutator).

After that it would be naturally to define a superanalogue sdet of the determinant and establish the identity

$$\operatorname{sdet}(\exp A) = \exp(\operatorname{str} A) \qquad (39)$$

which is well-known in ordinary case. It can be done, indeed, but not so straightforward as we acted until now.

The point is that to study the properties of determinants is natural on the language of Lie group theory,[10] and this theory is essentially non-linear.

Try, for instance to define a notion of a supergroup so that the set $\mathrm{GL}(V)$ of invertible linear operators in a super vector space $V$ were an example of such object. The correct answer is not simple and is formulated on the language of the theory of supermanifolds (see the end of this section).

### 3.3 Supersymmetry in analysis

Assume that we have several even variables $x_1,\, x_2,\, \ldots,\, x_n$ and several odd variables $\xi_1,\, \xi_2,\, \ldots,\, \xi_m$. A function of all these variables is by definition the expression

$$f(x,\, \xi) = \sum_I f_I(x_1,\, x_2,\, \ldots,\, x_n)\xi_I \qquad (40)$$

where the summation is over all subsets $I = \{i_1,\, i_2,\, \ldots,\, i_k\},\ 1 \le i_1 < i_2 < \cdots < i_k \le m,\ \xi_I$ means the product $\xi_{i_1}\xi_{i_1}\cdots\xi_{i_1}$ and $f_I(x)$ are smooth

[10] A lie group is a group which is also a smooth manifold, so that group multiplication is a smooth map. Most of groups used in different applications are Lie groups.

functions of even variables. We assume that they are real valued and have compact supports (i.e. vanish outside a certain compact set). It is clear that expressions (40) form a real associative algebra with respect to ordinary addition and multiplication, taking into account the relations $\xi_i\xi_j + \xi_j\xi_i = 0$. We denote this algebra by $C_0^\infty(\mathbb{R}^{n,m})$ and consider its elements (40) as smooth functions on a *supermanifold* $\mathbb{R}^{n,m}$

It is known that main operations of analysis, differentiation and integration, can be defined in pure algebraic terms, using the algebra of smooth functions. This approach to analysis can be easily extended to supermanifolds.

Recall that a *derivation* of an associative algebra $\mathcal{A}$ is a map $\partial : \mathcal{A} \leftrightarrow \mathcal{A}$, satisfying the it Leibnitz rule:

$$\partial(ab) = \partial(a)b + a\partial(b).$$

**Exercise 30** *Show that any derivation of the algebra $C_0^\infty(\mathbb{R}^n)$ has the form*

$$\partial f = \sum_{k=1}^{n} a_k \partial_k f$$

*where $a_k \in C^\infty(\mathbb{R}^n)$ and $\partial_k = \partial/\partial x_k$ is the operator of the partial derivative with respect to $x_k$.*

For a superalgebra $\mathcal{A}$ it is natural to define a derivation as a map $\partial : \mathcal{A} \leftrightarrow \mathcal{A}$, satisfying the it super-Leibnitz rule:

$$\partial(ab) = \partial(a)b + (-1)^{p(a)p(\partial)} a\partial(b)$$

where $p(\partial)$ is a parity of the derivation in question. In the algebra $C_0^\infty(\mathbb{R}^{n,m})$ we can define the even derivations $\partial_k$ and odd derivations $\delta_k$ by the conditions

$$\partial_k x_j = \delta_{k,j}, \quad \partial_k \xi_j = 0, \quad \delta_k x_j = 0, \quad \delta_k \xi_j = \delta_{k,j}.$$

More general differential operators on a supermanifold $M$ can be also defined in terms of the function algebra $\mathcal{A}(M)$. I only recall how it can be done for ordinary manifolds. Denote by $M_f$ the operator of multiplication by a smooth function $f$ in $\mathcal{A}(M)$.

**Exercise 31** *Show that a map $D : \mathcal{A} \leftrightarrow \mathcal{A}$ is a differential operator of order $\leq k$ if and only if*

$$[\cdots[[D,\, M_{f_0}],\, M_{f_1}],\, \ldots,\, M_{f_k}] = 0$$

*for all $(k+1)$-tuples of smooth functions $f_0,\, f_1,\, \ldots,\, f_k$.*

**Exercise 32** *Prove that the algebra* $C^\infty(\mathbb{R}^{0,m})$ *is finite-dimensional and find its dimension.*

**Exercise 33** *Show that any linear operator* $A$ *in the space* $C^\infty(\mathbb{R}^{0,m})$ *is differential, i.e. has the form*

$$A = \sum_{I,J} a_{I,J} \xi_I \partial_J$$

*where* $I = \{i_1 < i_2 < \cdots < i_k\},\ J = \{j_1 < j_2 < \cdots < j_l\}$ *and* $a_{I,J}$ *are real coefficients.*

Pass now to the integration over a superspace $\mathbb{R}^{n,m}$. The ordinary integral over $\mathbb{R}^n$ is determined up to a scalar factor by the properties:

1. Integral is a continuous linear functional on $C_0^\infty(\mathbb{R}^{n,m})$.
2. Images of operators $\partial_i$, $i \leq i \leq n$, are contained in the kernel of this functional.

It is natural to define the integral over a superspace so that it possessed the analogous properties.

Therefore, we come to the formula

$$\int_{\mathbb{R}^{n,m}} f(x, \xi) d^n x\, d^m \xi = \int_{\mathbb{R}^n} f_{12\ldots m}(x) d^n x. \tag{41}$$

This very formula was in the source of integral calculus on supermanifolds.

**Exercise 34** *Show that any linear operator in the space* $C^\infty(\mathbb{R}^{0,m})$ *can be written as an integral operator:*

$$(Af)(\xi) = \int_{\mathbb{R}^{0,m}} A(\xi, \eta) f(\eta) d^m \eta \tag{42}$$

*where* $A(\xi, \eta)$ *is a function of 2m odd variables and the integral over variables* $\eta$ *is supposed to be interchangeable with left multiplication by variables* $\xi$*.*

**Exercise 35** *Prove that the trace of the integral operator* (16) *can be computed by the formula*

$$\operatorname{tr} A = \int_{\mathbb{R}^{0,m}} A^\vee(\xi, \xi) d^m \xi \tag{43}$$

*where the operation sends a monomial* $\xi_I \eta_J$ *to the monomial* $\eta_J \xi_I$*.*

A very important property of an integral is its behavior under the change of variables. For an ordinary (even) integral this behavior is essentially encoded in the notation (in the case of one variable this notation was suggested by Leibnitz): the integrand is not a function $f(x)$, but the *differential form* $f(x)d^nx$. Here $d^nx$ is a short notation for the $n$-form $dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n$. If the variables $x^i$ are the functions of another variables $y^j$, then $d^nx$ transforms into $\frac{D(x)}{D(y)}d^ny$, where $fracD(x)D(y)$ is the Jacobian of the change of variables, equal to the determinant of the Jacobi matrix $\left\| \begin{array}{ccc} \partial x^1/\partial y^1 & \cdots & \partial x^1/\partial y^n \\ \cdots & \cdots & \cdots \\ \partial x^n/\partial y^1 & \cdots & \partial x^n/\partial y^n \end{array} \right\|$.

Note that this rather cumbersome rule is an immediatecorollary of the two simple principles:

1. The rule of the differential:

$$dx^i = \sum_j \partial x^i/\partial y^j dy^j.$$

2. The properties (associativity and anticommutativity) of the wedge product:

$$(a \wedge b) \wedge c = a \wedge (b \wedge c), \qquad a \wedge b = -b \wedge a.$$

In the odd case the expression $d^m\xi$ is no longer a tensor! It is a new geometric object which is called an *integral* form, or *Berezin* form in the honor of one of the founders of supersymmetry F.A. Berezin (1931 - 1980). The general definition of integral forms can be done in the context of the theory of supermanifolds. Here I mention only that already under linear changes of variables the quantity $d^m\xi$, contrary to the quantity $d^nx$, is not multiplied, but is divided by the determinant of the corresponding matrix. One can see it on the following simple

**Example 10** *Let* $\xi \mapsto \eta_i = a_i \cdot \eta_i$, $1 \leq i \leq m$. *Then*

$$\int \xi_1\xi_2\cdots\xi_m d^m\xi = \int \eta_1\eta_2\cdots\eta_m d^m\eta = \prod_{i=1}^m a_i \cdot \int \xi_1\xi_2\cdots\xi_m d^m\eta,$$

*which implies* $d^m\eta = \left(\prod_{i=1}^m a_i\right)^{-1} d^m\xi$.

This fact has a crucial value in the quantum field theory, where it is used to fight with divergences. To those, who want to know more about this, my advise is to read the Preface and Introduction to the [14] and after that the materials quoted there (e.g. [17]).

**Exercise 36** *a) Compute the integral*

$$\int_{\mathbb{R}^{0,2m}} \exp\Big(\sum_{i,j} a_{i,j}\xi_i\xi_j\Big)$$

*where $A = \|a_{i,j}\|$ is a skew-symmetric nondegenerate matrix of order $2m$. Compare it with the even integral*

$$\int_{\mathbb{R}^n} \exp -\pi \sum_{i,j} b_{i,j}x^i x^j d^n x$$

*where $B = \|b_{i,j}\|$ is a real symmetric positively defined matrix.*

*b) Show that* det $A$ *of a $2m \times 2m$ skew-symmetric matrix is a full square of a certain polynomial $Pf(A)$ of degree $m$ in the matrix elements $a_{i,j}$.*

## 3.4 Supersymmetry and geometry

Let now pass to the geometry. As we saw above, both, algebra and analysis suggest to introduce the notion of a supermanifold. I shall not give here the rigorous definition, referring to the books [14, 13, 18]. Instead, I compare the definition of a supermanifold with the definition of an algebraic manifold in its modern form.

A naive definition of an algebraic manifold as a set of solutions of a system of algebraic equations in affine or projective space in modern mathematics is superseded by a functorial definition. Namely, an algebraic manifold $M$ over a field $K$ is considered as a functor from the category of commutative $K$-algebras to the category of sets. To each $K$-algebra $A$ there corresponds the set $M_A$ of $A$-points of $M$, i.e. solutions of the corresponding system of equations with coordinates from $A$.

If now we replace in this definition the term “commutative $K$-algebra” by ”supercommutative $\mathbb{Z}_2$-graded $K$-algebra”, we get the definition of a supermanifold.

Thus, we give up the set-theoretic approach and describe all properties of a supermanifold in terms of functions on it. In particular, instead of a map $M \leftrightarrow N$ we consider the corresponding homomorphism of $K$-algebras $C^\infty(N) \leftrightarrow C^\infty(M)$. The role of a point $m \in M$ is played by a maximal ideal in $C^\infty(M)$. Note, that a supermanifold, contrary to the classical case, *is not* determined by the set of its points.

Here is a simple but rather important example of a supermanifold. Consider the algebra $A$ of smooth functions $f(t, \tau)$ of one even variable $t$ and one odd variable $\tau$. We consider $A$ as an algebra of smooth functions on a supermanifold $M$ of dimension (1, 1) which is a superanalogue of a real line. In complete analogy with the even case, call a *vector field* on $M$ any derivation of the algebra $A$.

**Exercise 37** *Prove that any even (respectively, odd) vector field on $M$ has the form*

$$v = f(t)\frac{\partial}{\partial t} + g(t)\tau\frac{\partial}{\partial \tau} \qquad \left(\textit{resp.}\, \xi = \varphi(t)\tau\frac{\partial}{\partial t} + \psi(t)\frac{\partial}{\partial \tau}\right) \tag{44}$$

*where $f$, $g$, $\varphi, \psi$ are smooth real functions of $t$.*

The set of all vector fields on $M$ form a Lie superalgebra with respect to the operation of supercommutator:

$$[v_1,\, v_2] = v_1 v_2 - (-1)^{p(v_1)p(v_2)} v_2 v_1. \tag{45}$$

It is remarkable that an even field $v_0 = \frac{\partial}{\partial t}$ is a square of the odd field $\xi_0 = \frac{\partial}{\partial \tau} + \tau\frac{\partial}{\partial t}$: $v_0 = [\xi_0,\, \xi_0]$. This have the principal importance for the theoretical physics. Indeed, let the variable $t$ denote the physical time. All the evolution laws are formulated in terms of the field $v_0$. Hence, they are corollaries of more fundamental laws, formulated in terms of the field $\xi_0$.

Changing with time is a special case of the action of a Lie group on a manifold. In this case the group in question is the additive group of the real field $\mathbb{R}$ and the manifold is the phase space of a physical system. Another examples arise when our system possesses certain kind of symmetry. Until recently, only ordinary Lie groups are considered in the role of a symmetry group (translation group, rotation group, Lorentz and Poincare groups etc). The idea of supersymmetry is that Lie supergroups (i.e. group objects in the category of supermanifolds) also can play such a role. It turns out, that the most of ordinary symmetry groups can be naturally extended to the supergroups.

**Example 11** *Let $V = V_{\bar{0}} \oplus V_{\bar{1}}$ be a $\mathbb{Z}_2$-graded real vector space. Define a supergroup $GL(V)$ as follows. For any supercommutative associative superalgebra $\mathcal{A}$ we define $GL(V)_{\mathcal{A}}$ as the set of all invertible matrices $A$ of the form* (35) *such that even blocks have even elements and odd blocks have odd elements. Thus we defined $GL(V)$ as supermanifold. To define a supergroup structure we must construct morphisms*

$$GL(V) \times GL(V) \longrightarrow GL(V) \quad \textit{and} \quad GL(V) \longrightarrow GL(V),$$

*corresponding to multiplication law and the inversion map. We can't do it here because the morphisms of supermanifolds were not yet defined.*

**Exercise 38** *Show that the formula*

$$sdet\, A = \det\,(A_{00} - A_{01}A_{11}^{-1}A_{10}) \cdot \det\,(A_{11}^{-1}) \tag{46}$$

*defines a homomorphism of the group $GL(V)_{\mathcal{A}}$ to the multiplicative group $\mathcal{A}^{\times}$ of invertible elements of $\mathcal{A}$.*

This homomorphism is called *superdeterminant*, or *Berezinian* of the operator $A$. Actually, it defines the homomorphism of $GL(\mathbb{R}^{n,m})$ into $GL(\mathbb{R}^{1,1})$.

**Exercise 39** *Formulate the rule of changing variables in the integral* (41)*.*

**Exercise 40** *Prove the equality* (39)*.*

### 3.5 Supersymmetry and arithmetic

This is a new subject, and not so many papers are appeared in this domain. Here we follow the article by D.Spector "Supersymmetry and the Möbius function" in Comm. Math. Phys., 1980. The point is a physical interpretation of some number-theoretical functions and, in particular, the *Möbius function*:

$$\mu(n) = \begin{cases} (-1)^k & \text{if } n \text{ is a product of } k \text{ different primes} \\ 0 & \quad\text{otherwise.} \end{cases}$$

The main subject of study in the statistical mechanics is the so-called partition function Z as a function of the temperature of the system. It encodes the statistical properties of a system in thermodynamic equilibrium. In quantum theory this quantity is equal to $\operatorname{tr} e^{\beta P}$ where $P$ is the energy operator in a Hilbert space $H$ and $\beta$ is a parameter which is proportional to the inverse absolute temperature: $\beta = \frac{1}{kT}$, $k$ is the Boltzmann's constant. When the system possesses a supersymmetry, the operator $P$ has the form of $Q^2$ where $Q$ is odd skew-Hermitian operator. In this situation the Witten's formula holds:

$$\operatorname{str} e^{-\beta P} = \operatorname{ind}\ker P \tag{47}$$

where ker means the kernel of the operator $P$ and ind is the difference of dimensions of even and odd components (see \$ 1). In particular, the left hand side actually does not depend on $\beta$! Passing to the limit $\beta \to 0$ or $\beta \to \infty$ in (47), one can obtain several important relations depending on the system in question.

In the D. Spector's article a model is considered which contains bosonic and fermionic particles,labelled by prime numbers. Every pure state (a vector from an orthonormal basis in $H$) is determined by the total quantity of particle of each sort. Assume that there are $k_i$ bosonic particles and $\varepsilon_i$ fermionic particles of sort $p_i$. Here $k_i$ is a non-negative integer and $\varepsilon_i$ takes values 0, 1 due to the *Pauli principle.* This state of the system can be conveniently labelled by the pair $(N,\, d)$ where

$$N = \prod_i p_i^{k_i}, \qquad d = \prod_i p_i^{\varepsilon_i}.$$

The mathematical formulation of what physicists call Bose-Einstein (or Fermi -Dirac) statistics is as follows. Let $L_\pm$ denote the phase space of one-partical states (bosonic or fermionic). Then the full phase space has the form

$$H = S(L_+) \otimes \wedge(L_-)$$

where $S$ means the symmetric algebra and $\wedge$ denotes the exterior, or anti-symmetric algebra.

In our model $N$ can be any natural number and $d$ can be any divisor of $N$ which is squarefree (i.e. not divisible by any square). The parity of the state is determined by the number of factors in $d:\ (-1)^p(N,\, d) = \mu(d)$.

**Exercise 41** *Derive from the Witten formula the relation*

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \zeta(s)^{-1}. \tag{48}$$

Hints and answers.

# 4 Differential and integral calculus on a lattice

Elementary length – same as
fundamental length.

Physics Encyclopedic Dictionary.
Moscow, Soviet Encyclopedia, 1983.

## 4.1 Introduction

Physicists are still arguing about the structure of our space on very small scale: is it infinitely divisible or there exists a certain elementary length? In the latter case the role of real line must take some discrete (i.e. consisting of isolated points) set. Several arguments show that it is convenient to choose as such set an arithmetic progression with common difference $h$ or a geometric progression with common ratio $q$. These progressions can be finite or infinite from one or both sides. It turns out that on this "discrete line", or *lattice* there exists a beautiful analogue of the ordinary calculus.

## 4.2 Arithmetic progression

Let us start with an arithmetic progression and put for simplicity $h = 1$. So, instead of "continuous line" $\mathbb{R}$ we consider "discrete line" $\mathbb{Z}$. Let $f$ be a real valued function on $\mathbb{Z}$, i.e. infinite in both sides sequence of real numbers $\{f(k),\, k \in \mathbb{Z}\}$. Since $\mathbb{Z}$ is discrete, all such functions are continuous. Moreover, all functions are differentiable if we define the *lattice derivative* $\Delta f$ of the function $F$ by the formula

$$\Delta f(n) \,:=\, f(n+1) - f(n). \tag{49}$$

We also define the *lattice integral*

$$\int_m^N F(k)\Delta k \,:=\, \sum_{k=m}^{n-1} f(k). \tag{50}$$

From these definitions immediately follows the lattice analogue of the Newton-Leibniz formula: if $\Delta F = f$, then

$$\int_m^N F(k)\Delta k = F(n) - F(m). \tag{51}$$

Further, in a usual calculus the monomials

$$P_n(x) = \frac{x^n}{n!}.$$

play an important role. They are characterized by the properties:

$$P_n' = P_{n-1}, \quad P_n(0) = 0 \quad \text{for} \quad n > 0, \qquad P_0(x) \equiv 1. \tag{52}$$

The lattice analogue of $P_n(x)$ are the polynomials $\Pi_n(x)$ characterized by the conditions

$$\Delta\Pi_n = \Pi_{n-1}, \quad \Pi_n(0) = 0 \quad \text{for} \quad n > 0, \qquad \Pi_0(x) \equiv 1. \tag{53}$$

**Exercise 42** *Compute explicitly polynomials* $\Pi_n(x)$.

**Exercise 43** *Let* $f(x)$ *be a polynomial of degree* $\leq n$ *which takes integral values at all integral points.*

*a) Show that* $f(x)$ *is an integral linear combination of* $\Pi_k(x)$, $0 \leq k \leq n : f(x) = \sum_{k=0}^{n} m_k \Pi_k(x)$.

*b) Prove that in the formula above the coefficients* $m_k$ *can be computed as*

$$m_k = \Big(\Delta^k f\Big)(0).$$

This is a lattice analogue of the Taylor formula.

Using polynomials $\Pi_n$ one can derive many summation formulas. Consider, as an example, the sum $S_2(n) = \sum_{k=0}^{n} k^2$. We observe that $k^2 = 2\Pi_2(k) + \Pi_1(k)$. Therefore,

$$\begin{array}{c} S_2(n) = \int_{k=1}^{n+1} k^2 \Delta k = \int_{k=1}^{n+1} (2\Pi_2(k) + \Pi_1(k))\Delta k = \\ 2\Big(\Pi_3(n+1) - \Pi_3(1)\Big) + \Big(\Pi_2(n+1) - \Pi_2(1)\Big) = \\ \frac{1}{6} n(n+1)(2n+1). \end{array}$$

To obtain more general summation formulas one have to know, how to express polynomials $P_k$ in terms of $\Pi_k$ and vice versa. From the exercise above we see that for any $n$ both $\{P_k\}_{0 \leq k \leq n}$ and $\{\Pi_k\}_{0 \leq k \leq n}$ are bases in the space of all polynomials of degree $\leq n$. Therefore there exist the coefficients $\{a_{ni}\}$ and $\{b_{nj}\}$ such that

$$\Pi_n(x) = \sum_{i=0}^{n} a_{ni} P_i(x), \qquad P_n(x) = \sum_{j=0}^{n} b_{ni} \Pi_j(x). \tag{54}$$

**Exercise 44** *Prove that for $n \geq 1$*

$$a_{n0} = b_{n0} = 0, \qquad a_{n1} = \frac{(-1)^{n-1}}{n}, \qquad b_{n1} = \frac{1}{n!}.$$

To compute the remaining coefficients, we use the fact that the operator $\Delta$ can be expressed in the form

$$\Delta = e^D - 1, \quad \text{or, more generally} \quad \Delta_h = \frac{e^{hD} - 1}{h} \tag{55}$$

where $D = \frac{d}{dx}$ is the usual derivative and $h$ is the common difference for our arithmetic lattice.

Further, instead of computing every coefficient $\{a_{ni}\}$ and $\{b_{nj}\}$ separately, we compute all of them simultaneously, finding the *generating functions*

$$A(x, y) = \sum_{0 \leq k \leq n} a_{nk} x^n y^k, \qquad B(x, y) = \sum_{0 \leq k \leq n} b_{nk} x^n y^k.$$

**Exercise 45** *Prove that*

$$A(x, y) = \Big(1 - y(e^x - 1)\Big)^{-1}, \qquad B(x, y) = \Big(1 - y\log(1 + x)\Big)^{-1}.$$

Come back to summation formulas. We are interested mainly in in sums of type (50), i.e. lattice integrals. Since the formulas of lattice analysis in the limit $h \to 0$ turn to the formulas of ordinary calculus, it is natural to try to express the lattice integral in terms of ordinary antiderivative $F(x)$. Such an expression was found almost 3 centuries ago by Euler and Maclaurin. The idea of Euler was especially simple. Let $\mathcal{F}$ be a lattice antiderivative of $f$. It is related to the ordinary antiderivative $F$ as follows:

$$\Delta\mathcal{F} = f = Df, \quad \text{or} \quad \mathcal{F} = \frac{D}{e^D - 1} F.$$

Using Bernoulli numbers $B - k$ (see $ 2), we can write

$$\frac{x}{e^x - 1} = \sum_{n \geq 0} B_n \frac{x^n}{n!}, \quad \text{hence} \quad \mathcal{F} = \sum_{n \geq 0} B_n \frac{F^{(n)}}{n!}.$$

We come to the final formula

$$\sum_{k=m}^{n-1} f(k) = \sum_{s \geq 0} \frac{B_s}{s!} \Big(F^{(s)}(n) - F^{(s)}(m)\Big). \tag{56}$$

**Exercise 46** *Derive from the Euler-Maclorin formula the following symbolic equalities:*

$$S_k(n) = \frac{(B+n+1)^k - (B+1)^k}{k+1} \tag{57}$$

$$(B+1)K = B^k \quad \textit{for} \quad k \geq 2. \tag{58}$$

The meaning of these equalities is as follows: after expanding both sides in powers of $B$, replace each term $B^k$ by the number $B_k$. E.g., the second equality for $k = 2$ means $B_2 + 2B_1 + 1 = B_2$, which implies $B_1 = -\frac{1}{2}$.

**Exercise 47** *a) Compute first 10 Bernoulli numbers, using the* (58)*.*
*b) Compute the sum* $S_3(n)$*.*

Bernoulli numbers appear in analysis of trigonometric functions.

**Exercise 48** *Prove the following formulas:*

$$\cot x = \frac{1}{x} - \sum_{k\geq 1} |B_{2k}| \frac{2^{2k} x^{2k-1}}{(2k)!}; \qquad \tan x = \sum_{k\geq 1} |B_{2k}| \frac{2^{2k}(2^{2k}-1) x^{2k-1}}{(2k)!}.$$

Consider the lattice analogue of the exponential function $e^{\lambda x}$. In ordinary calculus we can define this function or by the series $\sum_{n\geq 0} \lambda^n P_n(x)$ or by differential equalion $f'(x) = \lambda f(x)$ with initial condition $f(0) = 1$.

**Exercise 49** *a) for which* $\lambda$ *the series* $E(x) := \sum_n \geq 0 \lambda^n \Pi_n(x)$ *is convergent?*
*b) show that* $E(x)$ *satisfies the difference equation* $(\Delta E)(x) = \lambda E(x)$*.*

## 4.3 Geometric progression

Here the variable $X$ runs through the set[11]

$$q^{\mathbb{Z}} = \{q^n,\, n \in \mathbb{Z}\}.$$

Let $D_q$ denote the difference derivative:

$$(D_q f)(x) = \frac{f(qx) - f(q)}{qx - q}. \tag{59}$$

[11]The symbol $q$ below has different meaning in different applications. It can be independent from $x$ variable, taking real or complex non-zero values; in other examples it is a number of elements of a finite field. There are interesting theories where $q$ is a root of unity. Finally, one can consider $q$ as a formal variable and consider polynomials, rational functions or formal Laurent series in $q$.

Continuing the analogy with the previous section, we define the family of polynomials $\{Q_n(x)\}_{n\geq 0}$ by the conditions:

$$D_q Q_n = Q_{n-1}, \quad Q_n(1) = 0 \quad \text{for} \quad n \geq 1, \qquad Q_0(x) \equiv 1. \tag{60}$$

**Exercise 50** *Show that* $Q_n(x) = \frac{x^n}{(n_q)!}$

where $(n_q)! := 1_q \cdot 2_q \cdot \dots \cdot n_q \cdot$ and

$$n_q := 1 + q + q^2 + \dots + q^{n-1} = \frac{q^n - 1}{q - 1}.$$

So, polynomials $Q_n$ are obtained from $P_n$ by replacement of their denominators by their *quantum analogue*, or *q-analogue* $(n_q)!$, while the $\Pi_n$ were obtained from $P_n$ by replacement of their numerators $x^n$ by their lattice analogue $x(x-h)(x-2h)\cdots(x-(n-1)h)$.

Concider now the $q$-exponent function

$$\exp_q(\lambda x) := \sum_{n\geq 0} \lambda^n Q_n(x). \tag{61}$$

From this definition it is easy to derive that $\exp_q(x)$ satisfies to the difference equation

$$D_q \exp_q(\lambda x) = \lambda \cdot \exp_q(\lambda x) \tag{62}$$

which implies the identity

$$\exp_q(qx) = \exp_q(x) \cdot (1 + (q-1)x).$$

Replace in this identity $x$ by $qx$, $q^2 x$, ..., $q^{n-1}x$ and take the product. We get

$$\exp_q(q^n x) = \exp_q(x) \cdot \prod_{k=0}^{n-1} (1 + (q-1)q^k x).$$

Note, that in the algebra of formal power series in $q$ and in the algebra of analytic functions in $q, x$ for $|q| < 1$ the infinite product

$$\prod_{k=0}^{\infty} (1 + (q-1)q^k x)^{-1}$$

is convergent. Moreover, it satisfies to the same equation (62) and the same initial condition $f(0) = 1$.[12]

[12]This statement is not correct as it is, because our function is defined only on the set $q^{\mathbb{Z}}$, not containing zero. But this gap can be removed and we leave it to the reader.

Therefore,

$$\exp_q(x) = \prod_{k=0}^{n-1} (1 + (q-1)q^k x)^{-1}.$$

Replacing here $x$ by $\frac{x}{1-q}$, we come to a remarkable identity

$$\exp_q\left(\frac{x}{1-q}\right) = \prod_{k=0}^{n-1} (1 - q^k x)^{-1}.$$

Remembering the definition of $\exp_q(x)$, we come to the identity

$$\prod_{k=0}^{n-1} (1 - q^k x)^{-1} = \sum_{n \geq 0} \frac{x^n}{(1-q)(1-q^2)\cdots(1-q^{n-1})}. \tag{63}$$

In the present time for many elementary and special functions non-trivial $q$-analogues were found. They arise naturally in quantum theory, combinatorics, number theory and topology. I bring here three interpretations of the $q$-analogues of binomial coefficients

$$\begin{bmatrix} n \\ k \end{bmatrix}_q := \frac{(n_q)!}{(k_q)!\big((n-k)_q\big)!}.$$

1. If $q$ is the number of elements of a finite field $\mathbb{F}_q$, then $\begin{bmatrix} n \\ k \end{bmatrix}_q$ counts the number of $k$-dimensional subspaces in $n$-dimensional space over $\mathbb{F}_q$.

2. If $q$ is a complex number, $|q| = 1$ and $u$, $v$ are the coordinate functions of the quantum torus $\mathbb{T}_q$, related by $uv = qvu$, then

$$(u+v)^n = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_q u^k v^{n-k}.$$

3. Let $q$ be a formal variable. The expression $\begin{bmatrix} n \\ k \end{bmatrix}_q$ is a polynomial in $q$ of degree $k(n-k)$:

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \sum_{s=0}^{k(n-k)} m_{n,k}^s q^s.$$

It turns out that the coefficients $m_{n,k}^s$ have a remarkable representation-theoretic meaning. The group $S_n$ acts naturally in space $V_+ = \mathbb{C}[x_1, \ldots, x_n]$ of ordinary polynomials and in space $V_- = \wedge_{\mathbb{C}}\langle \xi_1, \ldots, \xi_n \rangle$ of polynomials in odd variables. Both spaces are graded by the degree of polynomials and both split into tensor products $V_\pm = I_\pm \otimes H_\pm$ of subspaces which consist respectively of invariant and harmonic polynomials. We have:

$$m_{n,k}^s = \dim \mathrm{Hom}_{S(n)}(H_+^s,\, H_-^k).$$

In the last time the basic facts about $q$-analogues of hypergeometric functions were intensively studied. See [6].

Hints and answers.

References (*out of date*)

# References


[1] Berezin F.A. *Method of second quantization*

[2] Bratelli O., Robinson D.W. *Operator algebras and quantum statistic mechanics* vol. 46 No 11 1999, 1999-1208

[3] Cartier P.

[4] *Encyclopaedia of Mathematical Sciences*, vol I-V, "VINITY", Moscow, 19??-2007, English translation: Springer,

[5] Grothendieck A.

[6] George Gasper and Mizan Rahman, *Basic Hypergeometric Series*, 2nd Edition, (2004), Encyclopedia of Mathematics and Its Applications, 96, Cambridge University Press, Cambridge. ISBN 0-521-83357-4.

[7] Husemoller D. *Fibre bundles*, Springer-Verlag, 1994.

[8] Kirillov A.A. *Elements of representation theory*, "Nauka", Moscow, 1972, 1978 (English transl. Springer, 1976.)

[9] Kirillov A.A., Gvishiani A.D. *Theorems and problems in Functional analysis*, "Nauka", Moscow, 1979, 1988 (English transl. Springer, 197?.)

[10] Kirillov A.A., Klumova I.N., Sossinski A.B. *Surreal numbers*, Kvant No , 1979.

[11] Knuth, D.E. Surreal Numbers, Addison-Wesley Publishing Company: Reading, Massachusetts. 1974.

[12] Koblitz N. *p-adic numbers, p-adic analysis and zeta-function*

[13] Leites D.A. *Introduction to supermanifolds*, Soviet Math. Uspekhi, vol. 35 (1980), 3-57

[14] Manin Yu.I. *Gauge fields and complex geometry*, "Nauka", Moscow, 1984. (English transl. ).

[15] Mandelbrot B. *The fractal geometry of Nature*, Freeman, San Francisco, 1982.

[16] Reshetikhin N.Yu., Tahtadjan L.A., Faddeev L.D. *Quantization of Lie groups and Lie algebras*, Algebra and Analysis vol. ? No ? (1989), 178-206.

[17] Slavnov A.A., Faddeev L.D., *Introduction to gauge field theory*, M. Nauka, 1978.

[18] Vladimirov V.S., Volovich I.V. *Introduction to superanalysis*, Teor. i Math. Physics, vol. 59, No 1, vol. 60, No 2 (1984), pages 3-27, 169-198

[19] Wikipedia: http://en.wikipedia.org/wiki/Portal:Mathematics