LECTURE 1
FUNCTIONS OF AN OPERATOR ARGUMENT

A.A. KIRILLOV

1. Introduction

Our goal is to make sense to the expressions like

\[ e^A, \quad \sin A, \quad \frac{A}{\sqrt{1 + A^2}}, \quad \text{or, in general,} \quad f(A) \]

where \( A \) is not a number but a linear operator in some vector space \( V \). We shall see that it can be done only under some restrictions on the function \( f \) and the operator \( A \) in question.

2. Functions of a matrix argument

2.1. Polynomial functions. We start with functions of the simplest kind – polynomial functions. Let \( P(x) = p_0 x^n + \cdots + p_{n-1} x + p_n \) be a polynomial with real or complex coefficients and \( A = || a_{i,j} || \) be a matrix of size \( N \). Then it is rather clear that the expression \( P(A) \) must be understood as the matrix

\[ P(A) = p_0 \cdot A^n + \cdots + p_{n-1} \cdot A + p_n \cdot 1 \]

where \( 1 \) denotes the unit matrix of size \( N \).

Remark 1. One can ask, why we should write the last summand in this form. The most natural answer is that only under this agreement we ensure the map \( P \mapsto P(A) \) to be a homomorphism of the polynomial algebra to the number field, i.e. the following equalities hold:

\[
(P_1 + P_2)(A) = P_1(A) + P_2(A), \quad (\lambda \cdot P)(A) = \lambda \cdot P(A),
\]

\[
(P_1 \cdot P_2)(A) = P_1(A) \cdot P_2(A)
\]

2.2. Rational functions. A rational function is by definition a ratio of two polynomial functions: \( R(x) = \frac{P(x)}{Q(x)} \). So, we can try to define the quantity \( R(A) \) as \( R(A) = \frac{P(A)}{Q(A)} \). But there is two delicate points. First, \( Q(A) \) must be invertible matrix; second, the matrix multiplication is not commutative, so we must choose between \( P(A) \cdot Q(A)^{-1} \) and \( Q(A)^{-1} \cdot P(A) \).

Actually, the second obstacle is inessential, because for a given matrix \( A \) all matrices of the form \( P(A) \) pairwise commute. So, \( P(A) \cdot Q(A) = Q(A) \cdot P(A) \), hence, \( Q(A)^{-1} \cdot P(A) = P(A) \cdot Q(A)^{-1} \).

Date: Sept 2007.
Consider the first point. For a given $A$ we can define $R(A)$ only for those $R = \frac{P(x)}{Q(x)}$, for which $Q(A)$ is invertible. Recall that a number $\lambda \in \mathbb{C}$ is called an eigenvalue of a matrix $A$ if $\det(A - \lambda \cdot \mathbf{1}) = 0$, i.e. when $(A - \lambda \cdot \mathbf{1})$ is not invertible. Collection of all eigenvalues of $A$ is called spectrum of $A$. We denote it by $\text{Spec } A$.

Proposition 1. The matrix $Q(A)$ is invertible if and only if $Q(\lambda) \neq 0$ for every eigenvalue $\lambda \in \text{Spec } A$.

Indeed, the polynomial $Q$ can be written as a product of linear factors:

$$Q(x) = c(x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_N)$$

where $\lambda_1, \lambda_2, \lambda_N$ are roots of $Q$ (taken with multiplicities). Therefore, $Q(A) = c(A - \lambda_1)(A - \lambda_2) \cdots (A - \lambda_N)$ and $Q(A)^{-1} = c^{-1}(A - \lambda_1 \cdot \mathbf{1})^{-1}(A - \lambda_2 \cdot \mathbf{1})^{-1} \cdots (A - \lambda_N \cdot \mathbf{1})^{-1}$.

So, $Q(A)$ is invertible when all matrices $(A - \lambda_i \cdot \mathbf{1})$ are. But this is the case when no roots of $Q$ belong to $\text{Spec } A$.

2.3. General functions. The statement of proposition 1 suggests that properties of the matrix $f(A)$ depend on behavior of the function $f$ on the spectrum of $A$. It is not completely true, but becomes true if we replace the set $\text{Spec } A$ by its infinitesimal neighborhood. To explain this, we start with two simple examples.

Example 1. Assume that our matrix $A$ is diagonal, or, more generally, can be reduced to the diagonal form by the transformation $A \mapsto SAS^{-1}$ with some invertible matrix $S$. Then for any polynomial function $F$ (hence, for any admissible rational function $F$) we have

$$F\left(\begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_N \end{pmatrix}\right) = \begin{pmatrix} F(a_1) & 0 & \cdots & 0 \\ 0 & F(a_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & F(a_N) \end{pmatrix}$$

This equality confirms the suggestion above and can be easily extended to all functions $F$ of a complex variable $z$.

But the life is not so simple.

Example 2. Suppose that our matrix $A$ can not be reduced to the diagonal form. The simplest example of such matrix is the so-called Jordan block...
$J_N(\lambda)$ of size $N$ with an eigenvalue $\lambda$ given by

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

The direct computation\footnote{There are several ways to make this computation in a “smart” way, practically with no computations at all. But it is very instructive to make it directly at least for a monomial $F(x) = x^n$.} shows that the answer has the following beautiful form:

$$F(J_N(\lambda)) = \begin{pmatrix}
F(\lambda) & F'(\lambda) & \frac{1}{2}F''(\lambda) & \ldots & \frac{1}{(N-1)!}F^{(N-1)}(\lambda) \\
0 & F(\lambda) & F'(\lambda) & \ldots & \frac{1}{(N-2)!}F^{(N-2)}(\lambda) \\
0 & 0 & F(\lambda) & \ldots & \frac{1}{(N-3)!}F^{(N-3)}(\lambda) \\
\ldots & \ldots & \ldots & \ldots \\
0 & 0 & 0 & \ldots & F'(\lambda) \\
0 & 0 & 0 & \ldots & F(\lambda)
\end{pmatrix}$$

We see, that the result depends not only on the values of $F$ on the spectrum of $A$ but also on the values of the first $N-1$ derivatives of $F$ at the points of Spec $A$. This is the exact meaning of the expression: “$F(A)$ depends on the values of $F$ on $U_{N-1}(\text{Spec } A)$ – the infinitesimal neighborhood of Spec $A$ of order $N - 1$”.

Introduce the notation

$$F_1 \sim F_2 \quad \text{on the set } X$$

which means that the difference $F_1 - F_2$ vanishes at all points $x \in X$ with multiplicity $\geq N$. We can also express this fact, saying that $F_1$ and $F_2$ coincide on $U_N(X)$.

Note, that for a polynomial $F$ the matrix $F(J_N(\lambda))$ can be non-zero even when $\lambda$ is a root of $F$. To have a zero value at $J_N(\lambda)$, $F$ must have $\lambda$ as a root of multiplicity at least $N$. In this case we say that $F$ vanishes on the infinitesimal neighborhood $U_{N-1}(\lambda)$.

Now we can discuss the case of a general matrix $A$. It is known that any matrix is similar to the direct sum of Jordan blocks of arbitrary sizes and eigenvalues. In other words, $A$ is similar to a block-diagonal matrices with $K$ blocks of the form

$$J_{N_k}(\lambda_k), \quad k = 1, 2, \ldots, K, \quad \text{where } \sum_{k=1}^{K} N_k = N.$$
Then for any polynomial function $P$ the value $P(A)$ is similar to the direct sum of $K$ blocks of the form
\begin{equation}
P(J_{N_k}(\lambda_k)), \quad k = 1, 2, \ldots, K.
\end{equation}
These matrices depend only on the values of $P$ on the infinitesimal neighborhood of Spec$A$ of order $\bar{N} - 1$, where
\begin{equation}
\bar{N} := \max_{1 \leq k \leq K} N_k.
\end{equation}

**Exercise 1.** Show that for any finite set $X \subset \mathbb{R}$, any $N \in \mathbb{N}$ and any smooth function $F$ on $\mathbb{R}$ there exist a polynomial $P$ such that $P^N \sim F$ on $X$.

Let now $F$ be any smooth function on $\mathbb{R}$. Choose a polynomial function $P$ satisfying $P^{\bar{N}-1} F$ on Spec$A$. Then the value $P(A)$ is determined uniquely and does not depend on the choice of $P$.

So, we put by definition
\begin{equation}
F(A) := P(A) \quad \text{for any polynomial } P \text{ satisfying } P^{\bar{N}-1} F \text{ on Spec } A;
\end{equation}

**Exercise 2.** Let $A = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}$. Compute the following matrices
a) $\sin A$; b) $|A|$; c) $e^A$; d) $\log A$ where
\[
\log(re^{i\theta}) := \log r + i\theta \quad \text{for } r > 0 \text{ and } -\pi < \theta < \pi;
\]
e) $\log A$ where
\[
\log(re^{i\theta}) := \log r + i\theta \quad \text{for } r > 0 \text{ and } 0 < \theta < 2\pi.
\]