AN INTRODUCTION TO FUNCTIONS OF MATRICES

SCOTT YORK

ABSTRACT

The purpose of this work is to show that the idea of defining functions of real numbers by using convergent power series allows one to obtain well-defined functions of matrices. The main result shows that one can define functions of matrices for infinitely differentiable functions for which all the derivatives are exponentially bounded. In particular, one obtains well-defined functions of matrices for the sine, cosine an exponential functions.

Introduction

A classical method for defining functions of real numbers into the real numbers is obtained by considering convergent power series as follows: let $D := \{x \in \mathbb{R} : \sum_{n=0}^{\infty} a_n x^n \text{ is convergent}\}$. Then the function $f : D \subseteq \mathbb{R} \to \mathbb{R}$ given by $f(x) := \sum_{n=0}^{\infty} a_n x^n$ is well-defined because the domain D consist of all $x \in \mathbb{R}$ such that $f(x) = \sum_{n=0}^{\infty} a_n x^n \in \mathbb{R}$; i.e., for every $x \in D \subseteq \mathbb{R}$ there exist $y \in \mathbb{R}$ such that $f(x) = y$. Thus, we obtain a function from $D \subseteq \mathbb{R}$ into $\mathbb{R}$ defined by the convergent power series. This method is intrinsically using two properties of the real numbers. (i) The concept of having convergent series is defined by using the concept of limits for a suitable sequence (given by the sequences of partial sums of the series); However in $\mathbb{R}$ the convergence of a sequence is related to the idea of distance as $\lim_{n\to\infty} x_n$ exists and equals x if and only if $\lim_{n\to\infty} |x_n - x| = 0$ where $|x|$ is the distance between x and 0 in $\mathbb{R}$. (ii) the real numbers is a complete set; i.e., every Cauchy sequence in $\mathbb{R}$ converges to an element in $\mathbb{R}$. Although this sounds like a triviality, it is no longer the case if one considers only rational numbers. For example, if x_n is the truncation of $\sqrt{2}$ to n decimal places then $x_n \in \mathbb{Q}$ for all $n \in \mathbb{N}$ but $\lim_{n\to\infty} x_n = \sqrt{2} \notin \mathbb{Q}$ and $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence. In order to use the same approach for matrices instead of real numbers one would like to define $f(A) := \sum_{n=0}^{\infty} a_n A^n$ if $f(x) = \sum_{n=0}^{\infty} a_n x^n$ provided $x \in D$. In order for this definition to make sense, the series $\sum_{n=0}^{\infty} a_n A^n$ has to converge to a matrix in order to define a function of a matrix A with range in the matrices. As convergence is related to distance, one needs to provide the vector space of matrices with a distance and also show that Cauchy sequences in the vector space of matrices converge to an element in the space of matrices. Furthermore, as the map given by $L : M_1(\mathbb{R}) \to \mathbb{R}$ defined by $L([a]) = a$ is a bijection, where $M_1(\mathbb{R})$ is the set of all matrices of order 1 by 1, then the idea of defining functions of matrices by using the concept of convergent power series is basically an attempt to extend the idea of defining functions of real numbers by using convergent power series into a bigger vector space.

The purpose of this work is to show precisely in what sense one can extend this definition from $\mathbb{R}$ (or $M_1(\mathbb{R})$) into $M_k(\mathbb{R})$ (the vector space of matrices of order k by k) and to identify a class of functions for which one has a power series representation. In the first section we introduce a distance for the vector space of matrices and show that $M_k(\mathbb{R})$ with that distance forms a complete vector space, i.e., every Cauchy sequence of matrices converge in $M_k(\mathbb{R})$. In the second section we show that the concept of power series can be used to define functions of matrices, and in the third section we look at Taylor's Theorem in order to find conditions under which a function can be written as a power series. Finally, our last section shows concrete examples of how to calculate functions of matrices for the exponential, sine and cosine functions when the matrix is either nilpotent or diagonalizable.

1. The Banach Space of Matrices

Definition 1.1. A field F is a nonempty set with two operations, addition $+ : F \times F \to F$ and multiplication $\cdot : F \times F \to F$, such that for all $a, b, c \in F$ the following are satisfied: (i) $a+b = b+a$, (ii) $(a+b)+c = a+(b+c)$, (iii) there is an additive identity element in F (usually) denoted by 0; i.e., there is an element $0 \in F$ such that $a + 0 = a$ for all $a \in F$, (iv) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$, (v) $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(b + c) \cdot a = b \cdot a + c \cdot a$, (vi) $a \cdot b = b \cdot a$, (vii) there exists $1 \in F$ such that $a \cdot 1 = 1 \cdot a = a$, (viii) for each $a \in F \backslash \{0\}$ there exists $b \in F$ such that $a \cdot b = b \cdot a = 1$.

Examples of fields are $\mathbb{Q}, \mathbb{R}$, and $\mathbb{C}$ with classical addition and multiplication.

Definition 1.2. A vector space V over a field F is a nonempty set with two operations, (vector) addition $+ : V \times V \to V$ and (scalar) multiplication $\cdot : F \times V \to V$, such that for all $v, u, w \in V$ the following properties are satisfied: (i) $u + v = v + u$, (ii) $u + (v + w) = (u + v) + w$, (iii) there is an object in 0 in V, called the zero vector for V, such that $0 + u = u + 0 = u$ for all u in V, (iv) for each u in V, there is an object $-u$ in V, such that $u + (-u) = (-u) + u = 0$, (v) $k(u+v) = ku+kv$ for all $k \in F$, (vi) $(k + l)u = ku + kl$ for all $k, l \in F$, (vii) $k(lu) = (kl)u$ for all $k, l \in F$, (viii) $1u = u$.

It follows from definition (1.1) that F is a vector space over itself. Other examples of vector spaces are $\mathbb{R}^k$ over $\mathbb{R}$, $\mathbb{C}^k$ over $\mathbb{C}$, the vector space of matrices of order $k \times k$ denoted by $M_k(\mathbb{R})$ over $\mathbb{R}$, and the matrices of order $k \times k$ denoted $M_k(\mathbb{C})$ over $\mathbb{C}$ with classical addition and scalar product of matrices.

Definition 1.3. A norm on a vector space V is a function $\|\cdot\|_V : V \to [0, \infty)$ that satisfies the following properties:

(1) $\|x\|_V = 0$ if and only if $x = 0$.
(2) $\|\alpha \cdot x\|_V = |\alpha| \|x\|_V$ for all $x \in V$ and $\alpha \in F$.
(3) $\|x + y\|_V \le \|x\|_V + \|y\|_V$ for all $x, y \in V$(Triangular inequality).

In addition, $(V, \|\cdot\|)$ is called a complete normed space, or a Banach space, if every Cauchy sequence converges to an element in V.

Classical examples of normed vector spaces are $(\mathbb{R}, |\cdot|)$ over the field $\mathbb{R}$, $(\mathbb{Q}, |\cdot|)$ over the field $\mathbb{Q}$, and $(\mathbb{C}, |\cdot|)$ over the field $\mathbb{C}$ where $|\cdot|$ denotes the absolute value. However, $(\mathbb{Q}, |\cdot|)$ is not a complete normed space. In order to see this, consider, for example, $x := \sqrt{p}$ where p is some prime number. Then, clearly $x \notin \mathbb{Q}$ and if x_n is given by the truncation of $\sqrt{p}$ to n decimal places, (for example $x_3 = 1.732$ if p=3) then $x_n \in \mathbb{Q}$ for all $n \in \mathbb{N}$, and $\{x_n\}_{n \in \mathbb{N}}$ is a Cauchy sequence. Furthermore, $\lim_{n\to\infty} x_n = \sqrt{p}$ because $\lim_{n\to\infty} |x_n - \sqrt{p}| = 0$; but $\sqrt{p} \notin \mathbb{Q}$, and therefore $\mathbb{Q}$ is not a complete normed space.
Other examples are given by the vector spaces $\mathbb{R}^k$ or $\mathbb{C}^k$ in which several norms can be defined. An important result of Real Analysis asserts that all norms are equivalent in $\mathbb{R}^k$ [3, p.38]; i.e., if $\|\cdot\|_1$ and $\|\cdot\|_2$ are two norms in $\mathbb{R}^k$, then there exist $c, C > 0$ such that $c\|x\|_1 \le \|x\|_2 \le C\|x\|_1$ for all $x \in \mathbb{R}^k$. Furthermore, $\mathbb{R}^k$ with any norm is complete; i.e., $\mathbb{R}^k$ is a Banach space [2, Thm 23].

Lemma 1.4. *If $(V, \|\cdot\|_V)$ is a complete normed vector space, then $\|\cdot\|_V : V \to [0, +\infty)$ is continuous.*

Proof. In order to see this recall that a function $\eta : V \to \mathbb{R}$ is continuous if and only if $x_n \to x$ implies that $\eta(x_n) \to \eta(x)$ (see [5, Thm 21.3]), or equivalently, if $\|x_n - x\|_V \to 0$ then $|\eta(x_n) - \eta(x)| \to 0$. Let $x_n \to x$ in V as $n \to \infty$, and if $\eta(x) := \|x\|_V$ then

$$|\eta(x_n) - \eta(x)| = |\|x_n\|_V - \|x\|_V| = \begin{cases} \|x_n\|_V - \|x\|_V \text{ if } \|x_n\|_V \ge \|x\|_V, \\ \|x\|_V - \|x_n\|_V \text{ if } \|x_n\|_V \le \|x\|_V. \end{cases} \tag{1.1}$$

From the triangular inequality property of Definition 1.3, one obtains that

$$\|x\|_V \le \|x - x_n\|_V + \|x_n\|_V \text{ or } \|x\|_V - \|x_n\|_V \le \|x_n - x\|_V.$$

Similarly $\|x_n\|_V - \|x\|_V \le \|x_n - x\|_V$. By (1.1) follows that

$$|\eta(x_n) - \eta(x)| \le \|x_n - x\|_V . \tag{1.2}$$

Thus, if $\|x_n - x\|_V \to 0$ as $n \to \infty$, then $\eta(x_n) \to \eta(x)$ in $\mathbb{R}$; i.e., $\eta : V \to [0, +\infty)$ is continuous where $\eta(x) = \|x\|$. □

Definition 1.5. Let $A \in M_k(\mathbb{R})$ and define

$$\|A\|_{M_k} := \sup_{\substack{x\in\mathbb{R}^k \\ \|x\|_{\mathbb{R}^k}=1}} \|Ax\|_{\mathbb{R}^k} , \tag{1.3}$$

where $\|\cdot\|_{\mathbb{R}^k}$ denotes any norm in $\mathbb{R}^k$.

Since $T : \mathbb{R}^k \to \mathbb{R}^k$ defined by $Tx := Ax$ is a linear transformation, then $T : x \to Ax$ is continuous. Furthermore, since $\{x \in \mathbb{R}^k : \|x\|_{\mathbb{R}^k} = 1\}$ is compact in $\mathbb{R}^k$, it follows that if $A \in M_k(\mathbb{R})$ then, by Lemma 1.4, $\|A\|_{M_k} < \infty$, as any continuous function attains its maximum on a compact set, and $\|A\|_{M_k} = \sup_{\substack{x\in\mathbb{R}^k \\ \|x\|_{\mathbb{R}^k}=1}} \eta \circ T(x)$, where $\eta(x) = \|x\|$ as in Lemma 1.4. Thus, $\|A\|$ is finite for every $A \in M_k(\mathbb{R})$; i.e., $\|\cdot\|_{M_k}$ is a function with domain $M_k(\mathbb{R})$ and range $[0, +\infty)$.

Remark 1.6. Usually $\|A\|_{M_k}$ is defined by $\|A\|_{M_k} = \sup_{x\in\mathbb{R}^k} \dfrac{\|Ax\|_{R^k}}{\|x\|_{\mathbb{R}^k}}$. However, if $x \in \mathbb{R}^k$, then

$$\frac{1}{\|x\|_{\mathbb{R}^k}} \|Ax\|_{\mathbb{R}^k} = \left\| \frac{1}{\|x\|_{\mathbb{R}^k}} Ax \right\|_{\mathbb{R}^k} = \left\| A \frac{x}{\|x\|_{\mathbb{R}^k}} \right\|_{\mathbb{R}^k} .$$

Therefore, $\sup_{x\in\mathbb{R}^k} \dfrac{\|Ax\|_{\mathbb{R}^k}}{\|x\|_{\mathbb{R}^k}} = \sup_{x\in\mathbb{R}^k} \left\| A \dfrac{x}{\|x\|_{\mathbb{R}^k}} \right\|_{\mathbb{R}^k} = \sup_{\substack{x\in\mathbb{R}^k \\ \|x\|_{\mathbb{R}^k}=1}} \|Ax\|_{\mathbb{R}^k} = \|A\|_{M_k(\mathbb{R})}$; i.e., the two definitions are equivalent.

Lemma 1.7. *The function* $\|\cdot\|_{M_k} : M_k(\mathbb{R}) \to [0, +\infty)$ *defined by* (1.3) *is a norm on* $M_k(\mathbb{R})$.

Proof. We only need to show that $\|\cdot\|_{M_k}$ satisfies Definition 1.3.

(1) If $\|A\|_{M_k} = 0$, then $\sup_{x\in\mathbb{R}^k} \dfrac{\|Ax\|_{\mathbb{R}_k}}{\|x\|_{\mathbb{R}_k}} = 0$ for all $x \in \mathbb{R}^k$. Therefore, $\|Ax\|_{\mathbb{R}^k} = 0$ for all $x \in \mathbb{R}^k$. Thus, $A = 0$. The reciprocal is obvious.

(2) $\|\alpha A\|_{M_k} = \sup_{\substack{x\in\mathbb{R}^k \\ \|x\|_{\mathbb{R}^k}=1}} \|\alpha Ax\|_{\mathbb{R}^k} = \sup_{\substack{x\in\mathbb{R}^k \\ \|x\|_{\mathbb{R}^k}=1}} |\alpha| \|Ax\|_{\mathbb{R}^k} = |\alpha| \sup_{\substack{x\in\mathbb{R}^k \\ \|x\|_{\mathbb{R}^k}=1}} \|Ax\|_{\mathbb{R}^k} = |\alpha| \|A\|_{M_k}$.

(3) Let A and B be matrices. Then $\|Ax + Bx\|_{\mathbb{R}^k} \le \|Ax\|_{\mathbb{R}^k} + \|Bx\|_{\mathbb{R}^k}$ for all $x \in \mathbb{R}^k$ and $\|A + B\|_{M_k} = \sup_{\substack{x\in\mathbb{R}^k \\ \|x\|_{\mathbb{R}^k}=1}} \|Ax + Bx\|_{\mathbb{R}^k} \le \sup_{\substack{x\in\mathbb{R}^k \\ \|x\|_{\mathbb{R}^k}=1}} \|Ax\|_{\mathbb{R}^k} + \sup_{\substack{x\in\mathbb{R}^k \\ \|x\|_{\mathbb{R}^k}=1}} \|Bx\|_{\mathbb{R}^k} = \|A\|_{M_k} + \|B\|_{M_k}$ by the properties of the supremum.

The lemma follows from (1),(2), and (3). □

Definition 1.8. A sequence $\{A_n\}_{n\in\mathbb{N}} \subseteq M_k(\mathbb{R})$ is said to converge to a matrix $A \in M_k(\mathbb{R})$ if $\lim_{n\to\infty} \|A_n - A\|_{M_k} = 0$; in which case, we denote this as $\lim_{n\to\infty} A_n = A$ or $A_n \to A$ as $n \to \infty$ in norm $\|\cdot\|_{M_k}$.

In the particular case that $\mathbb{R}^k$ is considered with the sum norm, i.e., $\|x\|_{\text{sum}} := \sum_{i=1}^{k} |x_i|$ then one denotes

$$\|A\|_{\text{s}} := \sup_{x\in\mathbb{R}^k} \frac{\|Ax\|_{\text{sum}}}{\|x\|_{\text{sum}}}. \tag{1.4}$$

Lemma 1.9. *If $A_n \to A$ in $\|\cdot\|_s$, then $A_n \to A$ in the $\|\cdot\|_{M^k}$.*

Proof. As any $\|\cdot\|_{\mathbb{R}^k}$ is equivalent to $\|\cdot\|_{\text{sum}}$, there exists $c, C > 0$ such that $c\|x\|_{\text{sum}} \leq \|x\|_{\mathbb{R}^k} \leq C\|x\|_{\text{sum}}$ for all $x \in \mathbb{R}^k$. Therefore, $\|Ax\|_{\mathbb{R}^k} \leq C\|Ax\|_{\text{sum}}$ and $c\|x\|_{\text{sum}} \leq \|x\|_{\mathbb{R}^k}$. Thus $\frac{\|Ax\|_{\mathbb{R}^k}}{\|x\|_{\mathbb{R}^k}} \leq \frac{C}{c}\frac{\|Ax\|_{\text{sum}}}{\|x\|_{\text{sum}}}$ which, by taking the supremum, implies that

$$\|A\|_{M_k} \leq \frac{C}{c}\|A\|_{\text{s}}. \tag{1.5}$$

Finally, if $\|A_n - A\|_{\text{s}} \to 0$ as $n \to \infty$, then $\|A_n - A\|_{M^k(\mathbb{R})} \to 0$ as $n \to \infty$, by (1.5), and the Lemma follows. ☐

Theorem 1.10. *$(M_k(\mathbb{R}), \|\cdot\|_s)$ is a complete normed vector space.*

Proof. Let $\{A_m\}_{m\in\mathbb{N}}$ be a Cauchy sequence in $M_k(\mathbb{R})$. If $x \in \mathbb{R}^k$ then $\{A_m x\}_{m\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}^k$, since

$$\left\|A_n\frac{x}{\|x\|} - A_m\frac{x}{\|x\|}\right\|_{\text{sum}} \leq \sup_{\substack{y\in\mathbb{R}^k \\ \|y\|_{\text{sum}}=1}} \|A_n y - A_m y\|_{\text{sum}} < \varepsilon$$

if $n, m > N$, where $N \in \mathbb{N}$ is such that $\|A_n - A_m\|_{\text{s}} < \varepsilon$ for $m, n > N$. As $\mathbb{R}^k$ is complete, there exists $y \in \mathbb{R}^k$ such that $A_n\frac{x}{\|x\|_{\text{sum}}} \to y$. Define $A\frac{x}{\|x\|_{\text{sum}}} := y$. It follows that $Ax = \|x\|_{\text{sum}}\, y$ and $A_n\frac{x}{\|x\|_{\text{sum}}} \to \frac{Ax}{\|x\|_{\text{sum}}}$. Therefore, $A_n x \to Ax$ as $n \to \infty$ in $\mathbb{R}^k$ with respect to the $\|\cdot\|_{\text{sum}}$ norm.

Now, we need to show that $A \in M_k(\mathbb{R})$; i.e.,

(1) $A(\alpha x) = \alpha Ax$ for all $x \in \mathbb{R}^k$.
(2) $A(x + z) = Ax + Az$ for all $x, z \in \mathbb{R}^k$.

However, (1) is obtained by noticing that $A\alpha x = \lim_{n\to\infty} A_n\alpha x = \alpha \lim_{n\to\infty} A_n x = \alpha Ax$. Similarly, one obtains (2). Thus, $A \in M_k(\mathbb{R})$.

Finally, we need to show that $A_n \to A$ in the matrix norm $\|\cdot\|_{\text{s}}$; i.e., for all $\varepsilon > 0$ there exist $N \in \mathbb{N}$ such that $\|A_n - A\|_{\text{s}} < \varepsilon$ for all $n > N$. Let $\varepsilon > 0$, since $\lim_{n\to\infty} A_n x = Ax$ for all $x \in \mathbb{R}^k$ with respect to $\|\cdot\|_{\text{sum}}$ norm, there exists $\{N_1, N_2 \ldots, N_k\} \subseteq \mathbb{N}$ such that $\|A_n e_i - Ae_i\|_{\text{sum}} < \varepsilon$ if $n > N_i$ where $\{e_1, \ldots, e_k\}$ denotes a basis for $\mathbb{R}^k$. It follows that

$$\begin{aligned}
\|A_n - A\| &= \sup_{\substack{x\in\mathbb{R}^k \\ \|x\|_{\text{sum}}=1}} \|A_n x - Ax\|_{\text{sum}} = \sup_{\substack{x\in\mathbb{R}^k \\ \|x\|_{\text{sum}}=1}} \left\|\sum_{i=1}^k x_i A_n e_i - \sum_{i=1}^k x_i Ae_i\right\|_{\text{sum}} \\
&= \sup_{\substack{x\in\mathbb{R}^k \\ \|x\|_{\text{sum}}=1}} \left\|\sum_{i=1}^k x_i(A_n e_i - Ae_i)\right\|_{\text{sum}} \\
&\leq \sup_{\substack{x\in\mathbb{R}^k \\ \|x\|_{\text{sum}}=1}} \sum_{i=1}^k |x_i|\,\|A_n e_i - Ae_i\|_{\text{sum}} \leq \varepsilon \sup_{\substack{x\in\mathbb{R}^k \\ \|x\|_{\text{sum}}=1}} \sum_{i=1}^k |x_i| \;=\; \varepsilon,
\end{aligned}$$

if $n > max\{N_1, \ldots, N_k\}$, and the result follows. ☐

Corollary 1.11. *$(M_k(\mathbb{R}), \|\cdot\|_{M_k})$ is a Banach space.*

Proof. If $\{A_n\}_{n\in\mathbb{N}}$ is Cauchy in $M_k(\mathbb{R})$ with the $\|\cdot\|_{M_k}$ norm, then $\{A_n x\}_{n\in\mathbb{N}}$ is Cauchy in $\mathbb{R}^k$ with the $\|\cdot\|_{\mathbb{R}^k}$ norm, but as all norms are equivalent to the sum norm in $\mathbb{R}^k$, $\{A_n x\}_{n\in\mathbb{N}}$ is Cauchy in $\mathbb{R}^k$ for the sum norm. By Theorem 1.10, $A_n \to A$ in the $\|\cdot\|_{\text{s}}$ norm; and, by Lemma 1.4, one obtains that $A_n \to A$ in the M_k norm as $n \to \infty$. ☐

2. Functions of Matrices Defined by Convergent Power Series

This section shows that convergent power series can be used to obtain well-defined functions from the Banach space of matrices into itself.

Theorem 2.1. *If* $\lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|} = r$, *then the power series* $\sum_{n=0}^{\infty} a_n z^n$ *has radius of convergence* r. *i.e.,* $\sum_{n=0}^{\infty} a_n z^n$ *converges for* $|z| < r$.

For a proof of Theorem 2.1 see [1].

Definition 2.2. If $\sum_{n=0}^{\infty} a_n x^n$ converges for all $x \in \mathbb{R}$, then $f(x) := \sum_{n=0}^{\infty} a_n x^n$ is called an entire function.

Remark 2.3. If $\sum_{n=0}^{\infty} a_n z^n$ converges for all $|z| < r$; i.e., $\lim \frac{|a_n|}{|a_{n+1}|} = r$, then by Theorem 2.1 $\sum_{n=0}^{\infty} |a_n|\, z^n$ has radius of convergence given by $\lim_{n\to\infty} \frac{\big|\,|a_n|\,\big|}{\big|\,|a_{n+1}|\,\big|} = r$; i.e., $\sum_{n=0}^{\infty} |a_n|\, z^n$ converges for $|z| < r$.

Example 2.4. Consider the power series given by

$$\sum_{k=0}^{\infty} \frac{1}{k!} x^k = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots .$$

Then

$$\lim_{n\to\infty} \left| \frac{a_n}{a_{n+1}} \right| = \lim_{n\to\infty} \frac{\frac{1}{n!}}{\frac{1}{(n+1)!}} = \infty.$$

Thus, Theorem 2.1 asserts that the power series $\sum_{k=0}^{\infty} \frac{1}{k!} x^k$ converges for all $x \in \mathbb{R}$; i.e., $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) := \sum_{k=0}^{\infty} \frac{1}{k!} x^k$ is an entire function.

Example 2.5. Consider the power series given by

$$\sum_{i=0}^{\infty} \frac{(-1)^i}{(2i)!} x^{2i} = 1 - \frac{1x^2}{2!} + \frac{1x^4}{4!} - \frac{1x^6}{6!} \cdots ;$$

i.e., $a_n = \frac{(-1)^n}{(2n)!}$. Thus $|a_n| = \frac{1}{(2n)!}$, $a_{n+1} = \frac{(-1)^{n+1}}{(2(n+1))!}$, and $|a_{n+1}| = \frac{1}{(2n+2)!}$. It follows that $\lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|} = \lim_{n\to\infty} \frac{\frac{1}{(2n)!}}{\frac{1}{(2n+2)!}} = \lim_{n\to\infty} \frac{(2n+2)!}{(2n)!} = \lim_{n\to\infty} (2n+1)(2n+2) = \infty.$

Thus, Theorem 2.1 asserts that the power series $\sum_{i=0}^{\infty} \frac{(-1)^i}{(2i)!} x^{2i}$ converges for all $x \in \mathbb{R}$; i.e., $g : \mathbb{R} \to \mathbb{R}$ given by $g(x) := \sum_{i=0}^{\infty} \frac{(-1)^i}{(2i)!} x^{2i}$ is an entire function.

Example 2.6. Consider the power series given by

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} = \frac{1x^1}{1!} + \frac{-1x^3}{3!} + \frac{1x^5}{5!} \cdots .$$

It follows that

$$\lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|} = \lim_{n\to\infty} \frac{(2n+3)!}{(2n+1)!} = \lim_{n\to\infty} (2n+1)(2n+3) = \infty.$$

Thus, Theorem 2.1 asserts that the power series $\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}$ converges for all $x \in \mathbb{R}$; i.e., $h : \mathbb{R} \to \mathbb{R}$ given by $h(x) := \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}$ is an entire function.

Lemma 2.7. *If $A \in M_k(\mathbb{R})$ then $\|Ax\|_{\mathbb{R}^k} \leq \|A\|_{M_k} \|x\|_{\mathbb{R}^k}$ for all $x \in \mathbb{R}^k$.*

The proof of Lemma 2.7 is straight forward from Definition 1.3. Notice that Lemma 2.7 implies that

$$\left\|A^i x\right\|_{\mathbb{R}^k} \leq \|A\|_{M_k}^i \|x\|_{\mathbb{R}^k} \text{ for all } x \in \mathbb{R}^k \text{ and } i \in \mathbb{N}. \tag{2.1}$$

Let $f(x) := \sum_{n=0}^{\infty} a_n x^n$ where $\sum_{n=0}^{\infty} a_n x^n$ converge for all $|x| < r$. Define $s_n(A) := \sum_{i=0}^{n} a_i A^i$, where $A \in M_k(\mathbb{R})$. Then

$$\|s_n(A) - s_m(A)\|_{M^k} = \sup_{x\in\mathbb{R}^k} \frac{\|s_n(A)x - s_m(A)x\|_{\mathbb{R}^k}}{\|x\|_{\mathbb{R}^k}}, \tag{2.2}$$

where $n, m \in \mathbb{N}$. By the triangular inequality, Lemma 2.7 and equation (2.1) one obtain that, for all $x \in \mathbb{R}^k$,

$$\|s_n(A)x - s_m(A)x\|_{\mathbb{R}^k} = \left\|\sum_{i=m}^{n} a_i A^i x\right\|_{\mathbb{R}^k} \leq \sum_{i=m}^{n} |a_i| \left\|A^i x\right\|_{\mathbb{R}^k} \leq \sum_{i=m}^{n} |a_i| \|A\|_{M_k}^i \|x\|_{\mathbb{R}^k} . \tag{2.3}$$

Thus

$$\|s_n(A) - s_m(A)\|_{M_k} \leq \sum_{i=m}^{n} |a_i| \|A\|_{M_k}^i . \tag{2.4}$$

If $x := \|A\|_{M_k}$, then $f(\|A\|_{M_k}) = \sum_{n=0}^{\infty} a_n \|A\|_{M_k}^n$ is convergent in $\mathbb{R}$ provided $\|A\|_{M_k} < r$ and by Remark 2.3 one obtains that $\sum_{n=0}^{\infty} |a_n| \|A\|_{M_k}^n$ is convergent in $\mathbb{R}$ if $\|A\|_{M_k} < r$. Thus, $\left\{\sum_{i=0}^{n} |a_i| \|A\|_{M_k}^i\right\}_{n\in\mathbb{N}}$ is Cauchy; i.e., for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$\left|\sum_{i=0}^{n} |a_i| \|A\|_{M_k}^i - \sum_{i=0}^{m} |a_i| \|A\|_{M_k}^i\right| < \varepsilon$$

if $n, m > N$. It follows from equation (2.4) that $\|s_n(A) - s_m(A)\|_{M_k} < \varepsilon$ if $n, m > N$; i.e., $\{s_n(A)\}_{n\in\mathbb{N}}$ is Cauchy in $M_k(\mathbb{R})$ and by Theorem 1.11 $\{s_n(A)\}_{n\in\mathbb{N}}$ is convergent in $M_k(\mathbb{R})$; i.e., one obtains the following result.

Theorem 2.8. *If $\sum_{n=0}^{\infty} a_n x^n$ converges for $|x| < r$, then the function $f(A) := \sum_{n=0}^{\infty} a_n A^n \in M_k(\mathbb{R})$ for all $A \in M_k(\mathbb{R})$ such that $\|A\|_{M_k} < r$ is well-defined.*

3. Power Series Representation of Exponentially Bounded Functions

The main purpose of this section is to show that the classical theorem of Taylor provides a basic framework to identify functions that can be written as power series. Thus, it will enable us to show explicit examples of functions of matrices by using the results of section 2.

Definition 3.1. If $f \in C^q(a,b)$ where $a < b$, $q \in \mathbb{N}$, and $x_0 \in (a,b)$, then the Taylor polynomial of degree q of f at x_0 is defined by

$$P_q[f,x_0](x) := \sum_{j=0}^{q} \frac{f^{(j)}(x_0)(x-x_0)^j}{j!}, \tag{3.1}$$

where $C^q(a,b)$ denotes the vector space of q-times differentiable continuous functions on (a,b) with values in $\mathbb{R}$.

Example 3.2. If $f \in C^2(\mathbb{R}), x_0 = 0, f(x_0) = 1, f'(x_0) = 2$, and $f''(x_0) = 3$, then

$$P_2[f,0](x) = 1 + \frac{2(x-x_0)}{1} + \frac{3(x-x_0)^2}{2}.$$

Theorem 3.3. *[Taylor] Let $f \in C^{q+1}(a,b)$ and $x_0 \in (a,b)$. Then*

$$f(x) = \sum_{j=0}^{q} \frac{f^{(j)}(x_0)(x-x_0)^j}{j!} + \frac{1}{q!}\int_{x_0}^{x} f^{(q+1)}(t)(t-x_0)^q dt \quad \textit{for all} \quad x \in (a,b).$$

For a proof of Theorem 3.3 see [2, Thm. 15, p. 148].

Proposition 3.4. $\lim_{n\to\infty} \frac{x^n}{n!} = 0$ *for all* $x \in \mathbb{R}$.

Proof. By Example 2.4

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} \text{ converges for all } x \in \mathbb{R}. \tag{3.2}$$

Recall that

$$\text{if the series } \sum_{k=0}^{\infty} y_k \text{ converges in } \mathbb{R} \text{ then } \lim_{n\to\infty} y_k = 0. \tag{3.3}$$

In this way if $y_k := \frac{x^k}{k!}$ for a fixed $x \in \mathbb{R}$ then $\sum_{k=0}^{\infty} y_k$ converges and $\lim_{n\to\infty} \frac{x_n}{n!} = \lim_{n\to\infty} y_k = 0$. Finally in order to show the validity of statement (3.3), one notices that the sequence of partial sums of the sequence $\{y_k\}_{k\in\mathbb{N}}$ is convergent; i.e., for all $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that $\left|\sum_{k=0}^{m} y_k - \sum_{k=0}^{n} y_k\right| < \varepsilon$ for all $m, n > N$. Let $m = n+1$, then $\left|\sum_{k=0}^{n+1} y_k - \sum_{k=0}^{n} y_k\right| < \varepsilon$; i.e., $|a_{n+1}| < \varepsilon$ for all $n > N$. Thus, $y_k \to 0$ as $k \to \infty$. □

Corollary 3.5. *Let $f \in C^\infty(a,b)$ if there exist $M > 0$ such that $\left|f^{(q)}(t)\right| \leq M$ for all $q \in \mathbb{N}$ and $t \in (a,b)$, then*

$$\lim_{q\to\infty} \frac{1}{q!}\int_{x_0}^{x} f^{(q+1)}(t)(t-x_0)^q dt = 0.$$

Proof.

$$\begin{aligned}\frac{1}{q!}\left|\int_{x_0}^{x} f^{(q+1)}(t)(t-x_0)^q dt\right| &\leq \frac{1}{q!}\int_{x_0}^{x}\left|f^{(q+1)}(t)\right|\left|t-x_0\right|^q dt\\ &\leq \frac{M}{q!}\int_{x_0}^{x}\left|t-x_0\right|^q dt\\ &= M\frac{(\left|x-x_0\right|)^{q+1}}{(q+1)!} \to 0 \text{ as } q\to\infty, \text{ by Proposition 3.4.}\end{aligned}$$

□

Corollary 3.6. *Let $f \in C^\infty(a,b)$. If there exist $M > 0$ and $\omega \in \mathbb{R}$ such that $\left|f^{(q)}(t)\right| \leq Me^{\omega t}$ for all $t \in (a,b)$ and all $q \in \mathbb{N}$, then*

$$\lim_{q\to\infty}\frac{1}{q!}\int_{x_0}^{x} f^{(q+1)}(t)(t-x_0)^q dt = 0.$$

Proof.

$$\begin{aligned}\frac{1}{q!}\left|\int_{x_0}^{x} f^{(q+1)}(t)(t-x_0)^q dt\right| &\leq \frac{M}{q!}\int_{x_0}^{x} e^{\omega t}\left|t-x_0\right|^q dt\\ &\leq \frac{M(max\left\{e^{x_0\omega}, e^{x\omega}\right\})}{q!}\int_{x_0}^{x}\left|t-x_0\right|^q dt\\ &= \frac{M(max\left\{e^{x_0\omega}, e^{x\omega}\right\})\left|x-x_0\right|^{q+1}}{(q+1)!} \to 0 \text{ as } q\to\infty.\end{aligned}$$

□

By combining Theorem 3.3 and Corollary 3.6 one obtains the following result.

Theorem 3.7. *Let $f \in C^\infty(a,b)$. If there exists $M > 0$ and $\omega \in \mathbb{R}$ such that $\left|f^{(q)}(x)\right| \leq Me^{\omega x}$ for all $x \in (a,b)$ and all $q \in \mathbb{N}$, then*

$$f(x) = \sum_{j=0}^{\infty}\frac{f^{(j)}(x_0)}{j!}(x-x_0)^j \textit{ for all } x \in (a,b).$$

Example 3.8. Let $f(x) := e^x$, then $\left|f^{(q)}(x)\right| = \left|e^x\right| \leq Me^{\omega x}$ for all $q \in \mathbb{N}$ if $M = 1$ and $\omega = 1$. Then Theorem 3.7 with $x_0 := 0$ asserts that

$$e^x = \sum_{k=0}^{\infty}\frac{1}{k!}x^k \text{ for all } x \in \mathbb{R}.$$

i.e., the exponential function is an entire function.

Example 3.9. Let $f(x) := \cos(x)$. Then $\left|f^{(q)}(x)\right| \leq 1 = 1\cdot e^{0\cdot x}$ for all $q \in \mathbb{N}$ if $M = 1$ and $\omega = 0$. Then Theorem 3.7 with $x_0 := 0$ asserts that

$$\cos(x) = \sum_{i=0}^{\infty}\frac{(-1)^i}{(2i)!}x^{2i} \text{ for all } x \in \mathbb{R}.$$

i.e., the cosine function is an entire function.

Example 3.10. Let $f(x) := \sin(x)$. Then $\left|f^{(q)}(x)\right| \leq 1 = 1\cdot e^{0\cdot x}$ for all $q \in \mathbb{N}$ if $M = 1$ and $\omega = 0$. Then Theorem 3.7 with $x_0 := 0$ asserts that

$$\sin(x) = \sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)!}x^{2k+1} \text{ for all } x \in \mathbb{R}.$$

i.e., the sine function is an entire function.

4. Examples of Functions of Matrices

The main result of our work is presented in this section in which we show how to obtain explicit formulas for functions of diagonalizable and nilpotent matrices when considering the functions that satisfy Theorem 3.7.

Example 4.1. Let $A := I$, the identity matrix in $M_k(\mathbb{R})$. Then $A^n = I^n = I$ for all $n \in \mathbb{N}$. From Example 3.9 one obtains that

$$\cos(I) = \sum_{n=0}^{\infty} \frac{(-1)^i}{(2i)!} I^{2i} = \begin{bmatrix} \cos(1) & 0 & \cdots & 0 \\ 0 & \cos(1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \cos(1) \end{bmatrix}.$$

Similarly, $\cos(tI)=\begin{bmatrix} \cos(t) & 0 & \cdots & 0 \\ 0 & \cos(t) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \cos(t) \end{bmatrix}$ for all $t \in \mathbb{R}$.

Remark 4.2. One could think that if $A \in M_2(\mathbb{R})$ then $\cos\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)=\begin{bmatrix} \cos(a) & \cos(b) \\ \cos(c) & \cos(d) \end{bmatrix}$; however, if $A := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, then $A^2 = 0$ and

$$\cos(A) = \sum_{i=0}^{\infty} \frac{(-1)^i}{(2i)!} A^{2i} = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \neq \begin{bmatrix} \cos(0) & \cos(1) \\ \cos(0) & \cos(0) \end{bmatrix} = \begin{bmatrix} 1 & \cos(1) \\ 1 & 1 \end{bmatrix}.$$

Definition 4.3. A matrix $A \in M_k(\mathbb{R})$ is called nilpotent of order m if $A^m = 0$, and $A^{m-1} \neq 0$ for some $m \in \mathbb{N}$.

Proposition 4.4. *If A is nilpotent of order m and f satisfies Theorem 3.7 with $x_0 := 0$, then $f(A) := \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} A^n = \sum_{n=0}^{m} \frac{f^{(n)}(0)}{n!} A^n$ provided $\|A\|_{M_k} < r$; i.e., one obtains a closed form for calculating functions of matrices A when A is nilpotent of order m.*

Example 4.5. Let A be a nilpotent matrix of order m, then $A^k = 0$, for all $k \geq m + 1$ and

$$\begin{aligned} e^A &:= \sum_{n=0}^{\infty} \frac{1}{n!} A^n \\ &= 1 + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots + \frac{A^m}{m!} + \frac{A^{m+1}}{(m+1)!} + \cdots + \\ &= 1 + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots + \frac{A^m}{m!}; \end{aligned}$$

i.e., $e^A = \sum_{n=0}^{m} \frac{A^n}{n!}$.

Example 4.6. Let m be even, i.e., $m = 2j$ for some $j \in \mathbb{N}$ and let $A \in M_k(\mathbb{R})$ be nilpotent of order m. Then

$$\begin{aligned}
\cos(A) &:= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} A^{2n} \\
&= I - \frac{A^2}{2!} + \frac{A^4}{4!} + \cdots + \frac{(-1)^j}{(2j)!} A^{2j} + \frac{(-1)^{j+1}}{(2(j+1))!} A^{2(j+1)} + \cdots + \\
&= I - \frac{A^2}{2!} + \frac{A^4}{4!} + \cdots + \frac{(-1)^j}{(2j)!} A^{2j} \\
&= \sum_{n=0}^{j} \frac{(-1)^n}{(2n)!} A^{2n}.
\end{aligned}$$

Similarly one obtains that

$$\begin{aligned}
\sin(A) &:= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} A^{2n+1} \\
&= A - \frac{A^3}{3!} + \frac{A^5}{5!} + \cdots + \frac{(-1)^j}{(2j-1)!} A^{2j-1} + \frac{(-1)^{j+1}}{(2(j+1)-1)!} A^{2(j+1)-1} \\
&= A - \frac{A^3}{3!} + \frac{A^5}{5!} + \cdots + \frac{(-1)^j}{(2j-1)!} A^{2j-1} \\
&= \sum_{n=0}^{j-1} \frac{(-1)^n}{(2n+1)!} A^{(2n+1)}.
\end{aligned}$$

Now, let m be odd; i.e., $m = 2j + 1$ for some $j \in \mathbb{N}$. Then

$$\begin{aligned}
\cos(A) &:= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} A^{2n} \\
&= I - \frac{A^2}{2!} + \frac{A^4}{4!} + \cdots + \frac{(-1)^j}{(2j)!} A^{2j} + \frac{(-1)^{j+1}}{(2(j+1))!} A^{2(j+1)} + \cdots \\
&= I - \frac{A^2}{2!} + \frac{A^4}{4!} + \cdots + \frac{(-1)^j}{(2j)!} A^{2j} \\
&= \sum_{n=0}^{j} \frac{(-1)^n}{(2n)!} A^{2n}.
\end{aligned}$$

Similarly one obtains that

$$\begin{aligned}
\sin(A) &:= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} A^{(2n+1)} \\
&= A - \frac{A^3}{3!} + \frac{A^5}{5!} + \cdots + \frac{(-1)^j}{(2j+1)!} A^{2j+1} + \frac{(-1)^{j+1}}{(2(j+1)+1)!} A^{2(j+1)+1} + \cdots \\
&= A - \frac{A^3}{3!} + \frac{A^5}{5!} + \cdots + \frac{(-1)^j}{(2j+1)!} A^{2j+1} \\
&= \sum_{n=0}^{j} \frac{(-1)^n}{(2n+1)!} A^{2n+1}.
\end{aligned}$$

Definition 4.7. A diagonal matrix $D := (d_{ij})$ is a matrix such that $d_{ij} = 0$ if $i \neq j$. Furthermore, a matrix $A \in M_k(\mathbb{R})$ is said to be diagonalizable if there exist a basis for $\mathbb{R}^k$ under which the associated linear transformation $T : \mathbb{R}^k \to \mathbb{R}^k$ given by $Tx = Ax$ can be represented as a diagonal matrix. Recall the following result from linear algebra.

Theorem 4.8. *Let $A \in M_k(\mathbb{R})$. The matrix A is diagonalizable if and only if there exist $B, D \in M_k(\mathbb{R})$ such that $A = BDB^{-1}$ where D is a diagonal matrix in $M_k(\mathbb{R})$.*

See [4] for a proof of Theorem 4.8.

Lemma 4.9. *If $D \in M_k(\mathbb{R})$ is a diagonal matrix, $f(x) := \sum_{n=0}^{\infty} a_n x^n$ converges for $|x| < r$, and $\|D\|_{M_k} < r$, then*

$$f(D) = \begin{bmatrix} f(d_{11}) & 0 & 0 & \cdots & 0 \\ 0 & f(d_{22}) & 0 & \cdots & 0 \\ 0 & 0 & f(d_{33}) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & f(d_{kk}) \end{bmatrix}.$$

Proof. Notice that $D^n = \begin{bmatrix} d_{11}^n & 0 & \cdots & 0 \\ 0 & d_{22}^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{kk}^n \end{bmatrix}$ for all $n \in \mathbb{N}$. Thus,

$$\begin{aligned} f(D) &= \sum_{n=0}^{\infty} a_n D^n = \sum_{n=0}^{\infty} a_n \begin{bmatrix} d_{11}^n & 0 & \cdots & 0 \\ 0 & d_{22}^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{kk}^n \end{bmatrix} \\ &= \begin{bmatrix} \sum_{n=0}^{\infty} a_n d_{11}^n & 0 & \cdots & 0 \\ 0 & \sum_{n=0}^{\infty} a_n d_{22}^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{n=0}^{\infty} a_n d_{kk}^n \end{bmatrix} \\ &= \begin{bmatrix} f(d_{11}) & 0 & \cdots & 0 \\ 0 & f(d_{22}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(d_{kk}) \end{bmatrix}. \end{aligned}$$

□

Theorem 4.10. *Let A be a diagonalizable matrix and f be an entire function. Then there exist $B, D \in M_k(\mathbb{R})$ such that D is diagonal and*

$$f(A) = Bf(D)B^{-1}.$$

Proof.

$$\begin{aligned} f(A) &= f(BDB^{-1}) = \sum_{n=0}^{\infty} a_n (BDB^{-1})^n = \sum_{n=0}^{\infty} a_n BD^n B^{-1} \\ &= B\left[\sum_{n=0}^{\infty} a_n D^n\right] B^{-1} \\ &= B \begin{bmatrix} f(d_{11}) & 0 & \cdots & 0 \\ 0 & f(d_{22}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(d_{kk}) \end{bmatrix} B^{-1} \\ &= Bf(D)B^{-1}. \end{aligned}$$

□

Example 4.11. Let $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. It is easy to see that if $P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$ then the inverse $P^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$ and

$$A = P \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} P^{-1}.$$

By Theorem 4.10 one obtains that

$$e^A = P \begin{bmatrix} e^5 & 0 \\ 0 & e^3 \end{bmatrix} P^{-1} = \begin{bmatrix} 2e^5 - e^3 & e^5 - e^3 \\ -2e^5 - 2e^3 & -e^5 - 2e^3 \end{bmatrix}$$

$$\cos(A) = P \begin{bmatrix} \cos(5) & 0 \\ 0 & \cos(3) \end{bmatrix} P^{-1} = \begin{bmatrix} 2\cos(5) - \cos(3) & \cos(5) - \cos(3) \\ -2\cos(5) - 2\cos(3) & -\cos(5) - 2\cos(3) \end{bmatrix}$$

$$\sin(A) = P \begin{bmatrix} \sin(5) & 0 \\ 0 & \sin(3) \end{bmatrix} P^{-1} = \begin{bmatrix} 2\sin(5) - \sin(3) & \sin(5) - \sin(3) \\ -2\sin(5) - 2\sin(3) & -\sin(5) - 2\sin(3) \end{bmatrix}.$$

References

1. Lars V. Ahlfors, *Complex analysis: An introduction of the theory of analytic functions of one complex variable*, Second edition, McGraw-Hill Book Co., New York, 1966. MR 0188405 (32 #5844)
2. R. Creighton Buck, *Advanced calculus*, third ed., McGraw-Hill Book Co., New York, 1978, With the collaboration of Ellen F. Buck, International Series in Pure and Applied Mathematics. MR 0476931 (57 #16482)
3. Serge Lang, *Real and functional analysis*, third ed., Graduate Texts in Mathematics, vol. 142, Springer-Verlag, New York, 1993. MR 1216137 (94b:00005)
4. David C. Lay, *Linear algebra and its applications*, third ed., Addison-Wesley, 2002.
5. James R. Munkres, *Topology: a first course*, Prentice-Hall Inc., Englewood Cliffs, N.J., 1975. MR 0464128 (57 #4063)

Department of Mathematical Sciences, Tennessee State University, Nashville, TN 37209

E-mail address: msyork21@gmail.com