An Introduction to Complex Differentials and Complex Differentiability

复微分和复可微性的介绍

Raphael Hunger 慕尼黑工业大学

Technical Report
TUM-LNS-TR-07-06

2007

Contents

1. Introduction

This technical report gives a brief introduction to some elements of complex function theory. First, general definitions for complex differentiability and holomorphic functions are presented. Since non-analytic functions are not complex differentiable, the concept of differentials is explained both for complex-valued and real-valued mappings. Finally, multivariate differentials and Wirtinger derivatives are investigated.

本技术报告简要介绍了复函数理论的一些要素。 首先，提出了复可微性和全纯函数的一般定义。 由于非解析函数不可复微分，因此对复值映射和实值映射都解释了微分的概念。 最后，研究了多元微分和 Wirtinger 导数

2. Complex Differentiability and Holomorphic Functions

Complex differentiability is defined as follows, cf. [Schmieder, 1993, Palka, 1991]:

Definition 2.0.1. *Let $\mathbb{A} \subset \mathbb{C}$ be an open set. The function $f : \mathbb{A} \to \mathbb{C}$ is said to be (complex) differentiable at $z_0 \in \mathbb{A}$ if the limit*

$$\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0} \tag{2.1}$$

exists independent of the manner in which $z \to z_0$. This limit is then denoted by $f'(z_0) = \frac{\mathrm{d}f(z)}{\mathrm{d}z}\big|_{z=z_0}$ and is called the derivative of f with respect to z at the point z_0.

A similar expression for (2.1) known from real analysis reads as

$$\frac{\mathrm{d}f(z)}{\mathrm{d}z} = \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z}, \tag{2.2}$$

where $\Delta z \in \mathbb{C}$ now holds. Note that if f is differentiable at z_0 then f is continuous at z_0. An equivalent, but geometrically more illuminating way to define the derivative follows from the linear approximation of f in the local vicinity of z_0 [Palka, 1991].

Definition 2.0.2. *Let $\mathbb{A}$ be an open set. The function $f : \mathbb{A} \to \mathbb{C}$ is said to be (complex) differentiable at $z_0 \in \mathbb{A}$ if there exists a complex-valued scalar g such that*

$$f(z) = f(z_0) + g \cdot (z - z_0) + e(z, z_0), \tag{2.3}$$

holds for every $z \in \mathbb{A}$ and the function $e(\cdot, \cdot)$ satisfies the condition

$$\lim_{z \to z_0} \frac{e(z, z_0)}{z - z_0} = 0. \tag{2.4}$$

The remainder term $e(z, z_0)$ in (2.4) obviously is $o(|z - z_0|)$ for $z \to z_0$ and therefore $g \cdot (z - z_0)$ dominates $e(z, z_0)$ in the immediate vicinity of z_0 if $g \neq 0$. Close to z_0, the differentiable function $f(z)$ can linearly be approximated by $f(z_0) + f'(z_0)(z - z_0)$. The difference $z - z_0$ is rotated by $\angle f'(z_0)$, scaled by $|f'(z_0)|$ and afterwards shifted by $f(z_0)$.

The concept of a differentiability in a single point readily extends to differentiability in open sets.

Definition 2.0.3. *Let $\mathbb{U} \subseteq \mathbb{A}$ be a nonempty open set. The function $f : \mathbb{A} \to \mathbb{C}$ is called holomorphic (or analytic) in $\mathbb{U}$, if f is differentiable in z_0 for all $z_0 \in \mathbb{U}$. Moreover, if f is analytic in the complete open domain-set $\mathbb{A}$, f is a holomorphic (analytic) function.*

An interesting characteristic of a function f analytic in $\mathbb{U}$ is the fact that its derivative f' is analytic in $\mathbb{U}$ itself [Spiegel, 1974]. By induction, it can be shown that derivatives of all orders exist and are analytic in $\mathbb{U}$ which is in contrast to real-valued functions, where continuous derivatives need not be differentiable in general. However, basic properties for the derivative of a sum, product, and composition of two functions known from real-valued analysis remain inherently valid in the complex domain. Assume that $f(z)$ and $g(z)$ are differentiable at z_0. Then, the following propositions hold:

Proposition 2.0.1. *The sum $f + g$ is differentiable at z_0 and*

$$(f + g)'(z_0) = f'(z_0) + g'(z_0). \tag{2.5}$$

Proposition 2.0.2. *The product fg is differentiable at z_0 and*

$$(fg)'(z_0) = f'(z_0)g(z_0) + f(z_0)g'(z_0). \tag{2.6}$$

Proposition 2.0.3. *If $g(z_0) \neq 0$, the quotient $\frac{f}{g}$ is differentiable at z_0 and*

$$\left(\frac{f}{g}\right)'(z_0) = \frac{f'(z_0)g(z_0) - f(z_0)g'(z_0)}{g^2(z_0)}. \tag{2.7}$$

Proposition 2.0.4. *If f is differentiable at $g(z_0)$, the composition $f \circ g$ is differentiable at z_0 and*

$$(f \circ g)'(z_0) = f'(g(z_0))g'(z_0) \quad \text{(chain rule)}. \tag{2.8}$$

Complex differentiability is closely related to the *Cauchy-Riemann* equations [Lang, 1993]. A necessary condition for f being holomorphic in $\mathbb{U}$ requires the *Cauchy-Riemann* equations to be satisfied.

Theorem 2.0.1: *Let $f(z) = u(z) + \mathrm{j}v(z)$ with $u(z), v(z) \in \mathbb{R}$ and $z = x + \mathrm{j}y$ with $x, y \in \mathbb{R}$. In terms of x and y, the function $f(z)$ can be expressed as $F(x, y) = U(x, y) + \mathrm{j}V(x, y)$ with $U(x, y), V(x, y) \in \mathbb{R}$. A necessary condition for $f(z)$ being holomorphic in $\mathbb{U}$ is that the following system of partial differential equations termed Cauchy-Riemann-equations holds for every $z = x + \mathrm{j}y \in \mathbb{U}$:*

$$\frac{\partial U(x, y)}{\partial x} = \frac{\partial V(x, y)}{\partial y} \quad and \quad \frac{\partial U(x, y)}{\partial y} = -\frac{\partial V(x, y)}{\partial x}. \tag{2.9}$$

Proof: According to Definition 2.0.3, $f(z)$ is holomorphic in $\mathbb{U}$ if $f(z)$ is differentiable at every $z \in \mathbb{U}$. Differentiability at z implies that the limit

$$\lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z}\bigg|_{z = x + \mathrm{j}y} = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{F(x + \Delta x, y + \Delta y) - F(x, y)}{\Delta x + \mathrm{j}\Delta y}$$

exists no matter which curve Δz moves along when approaching zero, see Definition 2.0.1 and (2.2). Setting $\Delta z = \Delta x + \mathrm{j}\Delta y$, two possible curves for $\Delta z \to 0$ are considered. The first curve goes in the horizontal direction with $\Delta y = 0$ and $\Delta x \to 0$ yielding

$$\begin{aligned} f'(z = x + \mathrm{j}y) &= \lim_{\Delta x \to 0} \frac{F(x + \Delta x, y) - F(x, y)}{\Delta x} \\ &= \lim_{\Delta x \to 0} \left\{ \frac{U(x + \Delta x, y) - U(x, y)}{\Delta x} + \mathrm{j}\frac{V(x + \Delta x, y) - V(x, y)}{\Delta x} \right\} \\ &= \frac{\partial U(x, y)}{\partial x} + \mathrm{j}\frac{\partial V(x, y)}{\partial x}. \end{aligned}$$

The second curve goes in the vertical direction with $\Delta x = 0$ and $\Delta y \to 0$ yielding

$$\begin{aligned} f'(z = x + \mathrm{j}y) &= \lim_{\Delta y \to 0} \frac{F(x, y + \Delta y) - F(x, y)}{\mathrm{j}\Delta y} \\ &= \lim_{\Delta y \to 0} \left\{ \frac{U(x, y + \Delta y) - U(x, y)}{\mathrm{j}\Delta y} + \mathrm{j}\frac{V(x, y + \Delta y) - V(x, y)}{\mathrm{j}\Delta y} \right\} \\ &= \frac{\partial U(x, y)}{\mathrm{j}\partial y} + \frac{\partial V(x, y)}{\partial y}. \end{aligned}$$

As both expressions have to be the same for $f(z)$ being holomorphic, (2.9) immediately follows as a necessary condition.

□

The next theorem provides conditions under which the Cauchy-Riemann equations are sufficient for $f(z)$ being holomorphic.

Theorem 2.0.2: *If the partial derivatives of $U(x, y)$ and $V(x, y)$ with respect to x and y are continuous, the Cauchy-Riemann equations are sufficient for $f(z)$ being holomorphic.*

Proof: See [Spiegel, 1974].   □

In the following, we give examples for analytic functions and functions which are not analytic.

Examples for analytic functions:

- $f(z) = z^n$ $\qquad$ $f'(z) = nz^{n-1}$

- $f(z) = \dfrac{az + b}{cz + d}$ $\quad$ $f'(z) = \dfrac{ad - bc}{(cz + d)^2}$

- $f(z) = \ln(z)$ $\qquad$ $f'(z) = \dfrac{1}{z}$

- $f(z) = \exp(az)$ $\quad$ $f'(z) = a\exp(az)$

Examples for *non*-analytic functions:

- $f(z) = |z|^2$

- $f(z) = \Re\{z\}$

- $f(z) = \Im\{z\}$

- $f(z) = z^*$

3. Differentials of Analytic and Non-Analytic Functions

二元

The total differential of the bivariate function $F(x, y)$ associated to the univariate function $f(z)$ via $F(x, y) = U(x, y) + \mathrm{j}V(x, y) = f(z)|_{z=x+\mathrm{j}y}$ reads as [Henrici, 1974]

$$\mathrm{d}F = \frac{\partial F(x, y)}{\partial x}\,\mathrm{d}x + \frac{\partial F(x, y)}{\partial y}\,\mathrm{d}y. \tag{3.1}$$

Of course, differentiability of $F(x, y)$ with respect to x and y in the real sense has to be imposed for the existence of the differential $\mathrm{d}F$ in (3.1). This implies the differentiability of the real-valued functions $U(x, y)$ and $V(x, y)$ with respect to x and y. Rewriting (3.1) by means of $F(x, y) = U(x, y) + \mathrm{j}V(x, y)$ yields

$$\mathrm{d}F = \frac{\partial U(x, y)}{\partial x}\,\mathrm{d}x + \mathrm{j}\frac{\partial V(x, y)}{\partial x}\,\mathrm{d}x + \frac{\partial U(x, y)}{\partial y}\,\mathrm{d}y + \mathrm{j}\frac{\partial V(x, y)}{\partial y}\,\mathrm{d}y. \tag{3.2}$$

Making use of

$$\begin{aligned}
\mathrm{d}z &= \mathrm{d}x + \mathrm{j}\mathrm{d}y, \\
\mathrm{d}z^* &= \mathrm{d}x - \mathrm{j}\mathrm{d}y,
\end{aligned} \tag{3.3}$$

the two differentials $\mathrm{d}x$ and $\mathrm{d}y$ can be expressed via

$$\begin{aligned}
\mathrm{d}x &= \frac{1}{2}\left(\mathrm{d}z + \mathrm{d}z^*\right) \\
\mathrm{d}y &= \frac{1}{2\mathrm{j}}\left(\mathrm{d}z - \mathrm{d}z^*\right).
\end{aligned} \tag{3.4}$$

Inserting (3.4) into the differential expression $\mathrm{d}F$ in (3.1) and reordering the result leads to

$$\begin{aligned}
\mathrm{d}F = &\frac{1}{2}\left[\frac{\partial U(x, y)}{\partial x} + \frac{\partial V(x, y)}{\partial y} + \mathrm{j}\left(\frac{\partial V(x, y)}{\partial x} - \frac{\partial U(x, y)}{\partial y}\right)\right]\mathrm{d}z \\
&+ \frac{1}{2}\left[\frac{\partial U(x, y)}{\partial x} - \frac{\partial V(x, y)}{\partial y} + \mathrm{j}\left(\frac{\partial V(x, y)}{\partial x} + \frac{\partial U(x, y)}{\partial y}\right)\right]\mathrm{d}z^*.
\end{aligned} \tag{3.5}$$

A major result can already be anticipated here.

Proposition 3.0.1. *The differential of any analytical function $f(z)$ does not depend on the differential $\mathrm{d}z^*$.*

Proof: Since any analytical function $f(z)$ satisfies the *Cauchy-Riemann* equations in (2.9), the factor in front of $\mathrm{d}z^*$ in the second line of (3.5) is zero. Obviously, the differential $\mathrm{d}F$ does not depend on $\mathrm{d}z^*$.

<div align="right">□</div>

Note that the converse of Proposition 3.0.1 is also true: If the differential of a function f does not depend on $\mathrm{d}z^*$, the function f is analytical.

Rearranging the terms in (3.5), the differential $\mathrm{d}F$ can be expressed as

$$
\begin{aligned}
\mathrm{d}F = &\frac{1}{2}\left[\frac{\partial}{\partial x}\left(U(x,y)+\mathrm{j}V(x,y)\right) - \mathrm{j}\frac{\partial}{\partial y}\left(U(x,y)+\mathrm{j}V(x,y)\right)\right]\mathrm{d}z \\
&+ \frac{1}{2}\left[\frac{\partial}{\partial x}\left(U(x,y)+\mathrm{j}V(x,y)\right) + \mathrm{j}\frac{\partial}{\partial y}\left(U(x,y)+\mathrm{j}V(x,y)\right)\right]\mathrm{d}z^*.
\end{aligned}
$$

Recognizing that $U(x,y) + \mathrm{j}V(x,y) = F(x,y)$, we finally obtain by factoring out the partial differential operators

$$
\mathrm{d}F = \frac{1}{2}\left[\frac{\partial}{\partial x} - \mathrm{j}\frac{\partial}{\partial y}\right]F(x,y)\,\mathrm{d}z + \frac{1}{2}\left[\frac{\partial}{\partial x} + \mathrm{j}\frac{\partial}{\partial y}\right]F(x,y)\,\mathrm{d}z^*. \tag{3.6}
$$

According to the total differential for real-valued multivariate functions, the introduction of the two operators $\frac{\partial}{\partial z}$ and $\frac{\partial}{\partial z^*}$ is reasonable as it leads to the very nice description of the differential $\mathrm{d}f$, where the real-valued partial derivatives are hidden [Trapp, 1996].

Theorem 3.0.1: *The differential* $\mathrm{d}f$ *of a complex-valued function* $f(z) : \mathbb{A} \to \mathbb{C}$ *with* $\mathbb{A} \subset \mathbb{C}$ *can be expressed as*

$$
\mathrm{d}f = \frac{\partial f(z)}{\partial z}\,\mathrm{d}z + \frac{\partial f(z)}{\partial z^*}\,\mathrm{d}z^*. \tag{3.7}
$$

Proof: See the preceding derivation and the definition of the partial derivatives $\frac{\partial}{\partial z}$ and $\frac{\partial}{\partial z^*}$ in the following.

<div align="right">□</div>

Definition 3.0.1. *The two 'partial derivative' operators* $\frac{\partial}{\partial z}$ *and* $\frac{\partial}{\partial z^*}$ *are defined by*

$$
\begin{aligned}
\frac{\partial}{\partial z} &:= \frac{1}{2}\left[\frac{\partial}{\partial x} - \mathrm{j}\frac{\partial}{\partial y}\right], \\
\frac{\partial}{\partial z^*} &:= \frac{1}{2}\left[\frac{\partial}{\partial x} + \mathrm{j}\frac{\partial}{\partial y}\right],
\end{aligned} \tag{3.8}
$$

and are often referred to as the ==Wirtinger derivatives== *[Wirtinger, 1926] and correspond to half of the del-bar and del operator [Spiegel, 1974].*

Basic rules for this *Wirtinger* calculus are given in the next theorem.

Theorem 3.0.2: *For the Wirtinger derivatives, the common rules for differentiation known from real-valued analysis concerning the sum, product, and composition of two functions hold as well. In particular,*

$$\frac{\partial}{\partial z} z^* = \frac{\partial}{\partial z^*} z = 0,$$

which means that z^* can be regarded as a constant when differentiating with respect to z, as well as z can be regarded constant when differentiating with respect to z^*.

Proof: With $z = x + \mathrm{j}y$ and $z^* = x - \mathrm{j}y$, Theorem 3.0.2 follows immediately from (3.8).   □

Examples:

- $\dfrac{\partial}{\partial z}|z|^2 = \dfrac{\partial}{\partial z}\left(zz^*\right) = z^*$

- $\dfrac{\partial}{\partial z} \exp\left(-|z|^2\right) = \dfrac{\partial}{\partial z}\exp\left(-zz^*\right) = -z^*\exp\left(|z|^2\right)$

Corollary 3.0.1. *Derivatives of the conjugate function $f^*(z)$ satisfy the relationships*

$$\frac{\partial f^*(z)}{\partial z} = \left(\frac{\partial f(z)}{\partial z^*}\right)^* \quad and \quad \frac{\partial f^*(z)}{\partial z^*} = \left(\frac{\partial f(z)}{\partial z}\right)^*. \tag{3.10}$$

Proof: See the definition of the Wirtinger derivatives in (3.8).   □

4. Differentials of Real-Valued Functions

Optimizations in communications and signal processing are frequently targeted on the maximization of a utility or on the minimization of a cost. Hence, most objectives are real-valued, as the standard total order can only handle real-valued arguments. On account of this, this chapter deals with functions $f(z) : \mathbb{U} \rightarrow \mathbb{R}$ having complex-valued arguments $z \in \mathbb{U} \subset \mathbb{C}$ that are mapped to real-valued scalars $f(z) \in \mathbb{R}$. In addition, simplifications resulting from this circumstance are investigated.

First of all it is obvious that the only possibility of a real-valued function $f(z)$ with complex argument z for being analytic is that $f(z)$ is constant for all z of its domain. This follows from the Cauchy-Riemann equations in (2.9), since $v(z) = V(x, y) = \Im\{f(z)\} = 0$ for real-valued $f(z)$. This leads to the following proposition.

> **Proposition 4.0.1.** *All non-trivial (not constant) real-valued functions $f(z)$ mapping $z \in \mathbb{A} \subset \mathbb{C}$ onto $\mathbb{R}$ are non-analytic functions and therefore not complex differentiable.*

With the definition of the differential $\mathrm{d}F$ in (3.2) and (3.6), it is easy to prove the following theorem.

Theorem 4.0.1: *The differential $\mathrm{d}f$ of a real-valued function $f(z) : \mathbb{A} \rightarrow \mathbb{R}$ with complex-valued argument $z \in \mathbb{A} \subset \mathbb{C}$ can be expressed as*

$$\mathrm{d}f = 2\Re\left\{\frac{\partial f(z)}{\partial z}\,\mathrm{d}z\right\} = 2\Re\left\{\frac{\partial f(z)}{\partial z^*}\,\mathrm{d}z^*\right\} \tag{4.1}$$

and is equivalent to

$$\mathrm{d}F = \frac{\partial F(x, y)}{\partial x}\,\mathrm{d}x + \frac{\partial F(x, y)}{\partial y}\,\mathrm{d}y. \tag{4.2}$$

Due to the property that the non-trivial real-valued functions are not analytic, stationary points of the real-valued $f(z)$ cannot be obtained by searching for points z where the derivative $f'(z)$ is zero. However, we can detect stationary points z of $f(z)$ by a vanishing differential $\mathrm{d}f$.

Theorem 4.0.2: *The differential* $\mathrm{d}f$ *of a real-valued function* $f(z) : \mathbb{A} \to \mathbb{R}$ *with complex argument* $z \in \mathbb{A} \subset \mathbb{C}$ *vanishes if and only if the Wirtinger derivative is zero:*

$$\mathrm{d}f = 0 \Leftrightarrow \frac{\partial f(z)}{\partial z} = 0. \tag{4.3}$$

Proof: First, we prove that $\frac{\partial f(z)}{\partial z} = 0$ leads to $\mathrm{d}f = 0$. This is a result from (4.1). The converse is shown by the following reasoning. For arbitrary ratios of $\mathrm{d}x$ and $\mathrm{d}y$, $\mathrm{d}F$ from (4.2) and therefore $\mathrm{d}f$ can only vanish if both partial derivatives of $F(x, y)$ with respect to x and y are zero. With

$$\frac{\partial f(z)}{\partial z}\,\mathrm{d}z = \frac{1}{2}\left[\frac{\partial U(x, y)}{\partial x} - \mathrm{j}\frac{\partial U(x, y)}{\partial y}\right](\mathrm{d}x + \mathrm{j}\mathrm{d}y), \tag{4.4}$$

vanishing partial derivatives document the second way of the equivalence relation in Theorem 4.0.2 as $F(x, y) = U(x, y)$ for real-valued $f(z)$.

$\square$

Gradient-based iterative algorithms targeted on maximizing or minimizing an objective function can be constructed by optimizing $\mathrm{d}z$ in the differential expression (4.1).

Corollary 4.0.1. *The steepest ascent of a real-valued function* $f(z) : \mathbb{A} \to \mathbb{R}$ *with complex-valued argument* $z \in \mathbb{A} \subset \mathbb{C}$ *is obtained for*

$$\mathrm{d}z = \frac{\partial f(z)}{\partial z^*}\,\mathrm{d}s, \tag{4.5}$$

where $\mathrm{d}s$ *is a real-valued differential. Thus, the steepest ascent points to the direction of* $\frac{\partial f(z)}{\partial z^*}$.

Proof: As the differential $\mathrm{d}f$ of a real-valued function can be expressed as (see Equ. 4.1)

$$\mathrm{d}f = 2\Re\left\{\frac{\partial f(z)}{\partial z^*}\,\mathrm{d}z^*\right\},$$

$\mathrm{d}f$ is maximized for real-valued $\frac{\partial f(z)}{\partial z^*}\,\mathrm{d}z^*$ if the norm of $\mathrm{d}z$ is fixed. Hence, $\mathrm{d}z^*$ has to be a scaled version of the conjugate of $\frac{\partial f(z)}{\partial z^*}$. Equivalently, $\mathrm{d}z$ must be a scaled version of $\frac{\partial f(z)}{\partial z^*}$ and (4.5) immediately follows.

$\square$

An iterative implementation could therefore read as

$$z \leftarrow z + 2\frac{\partial f(z)}{\partial z^*}\,\mathrm{d}s,$$

where $\mathrm{d}s$ can be interpreted as the step-size. Notice that there is a factor 2 in front of the Wirtinger derivative which follows from (4.1)!

5. Derivatives of Functions of Several Complex Variables
导数

When switching to functions of several complex variables stacked in the column vector $\boldsymbol{z} = [z_1, \ldots, z_n]^{\mathrm{T}} \in \mathbb{C}^n$, we confine ourselves to mappings onto the one-dimensional complex domain $\mathbb{C}$. For them, the holomorphic-property is defined by the following two definitions which are equivalent [Krantz, 1992]:

Definition 5.0.1. *A function* $f(\boldsymbol{z}) : \mathbb{C}^n \supset \mathbb{A} \to \mathbb{C}$ *is said to be holomorphic if for each* $k \in \{1, \ldots, n\}$ *and each fixed* $z_1, \ldots, z_{k-1}, z_{k+1}, \ldots, z_n$ *the function*

$$w \mapsto f([z_1, \ldots, z_{k-1}, w, z_{k+1}, \ldots, z_n]^{\mathrm{T}})$$

is holomorphic according to the one-dimensional sense in Definition 2.0.3 on the set $\{w \in \mathbb{C} : [z_1, \ldots, z_{k-1}, w, z_{k+1}, \ldots, z_n]^{\mathrm{T}} \in \mathbb{A}\}$.

This is nothing else than that the function $f(\boldsymbol{z})$ has to be holomorphic in each variable $z_1, \ldots, z_n$. An equivalent definition reads as follows [Krantz, 1992]:

Definition 5.0.2. *A function* $f(\boldsymbol{z}) : \mathbb{A} \to \mathbb{C}$ *that is continuously differentiable in each variable* z_k, $k \in \{1, \ldots, n\}$ *is said to be holomorphic if the Cauchy-Riemann equations are satisfied in each variable separately.*

Although there are many differences between univariate and multivariate complex functions, the Wirtinger calculus easily extends to the case of several complex variables.

Theorem 5.0.1: *The differential* $\mathrm{d}f$ *of a multivariate complex-valued function* $f(\boldsymbol{z}) : \mathbb{A} \to \mathbb{C}$ *with* $\mathbb{A} \subset \mathbb{C}^n$ *can be expressed as*

$$\begin{aligned} \mathrm{d}f &= \sum_{k=1}^{n} \frac{\partial f(\boldsymbol{z})}{\partial z_k} \, \mathrm{d}z_k + \sum_{k=1}^{n} \frac{\partial f(\boldsymbol{z})}{\partial z_k^*} \, \mathrm{d}z_k^* \\ &= \frac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{z}^{\mathrm{T}}} \, \mathrm{d}\boldsymbol{z} + \frac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{z}^{\mathrm{H}}} \, \mathrm{d}\boldsymbol{z}^*. \end{aligned} \tag{5.2}$$

Note that the operators $\frac{\partial}{\partial \boldsymbol{z}}$ and $\frac{\partial}{\partial \boldsymbol{z}^*}$ read as

$$\frac{\partial}{\partial \boldsymbol{z}} = \left[\frac{\partial}{\partial z_1}, \ldots, \frac{\partial}{\partial z_n}\right]^{\mathrm{T}},$$
$$\frac{\partial}{\partial \boldsymbol{z}^*} = \left[\frac{\partial}{\partial z_1^*}, \ldots, \frac{\partial}{\partial z_n^*}\right]^{\mathrm{T}}.$$

(5.3)

If they are applied to a scalar field $f(\boldsymbol{z})$ they generate a column vector of dimension n and mimic the gradient operator for real-valued functions. Again, both the differential $\mathrm{d}\boldsymbol{z} = [\mathrm{d}z_1, \ldots, \mathrm{d}z_n]^{\mathrm{T}}$ and its conjugate $\mathrm{d}\boldsymbol{z}^*$ are required to express the differential $\mathrm{d}f$ in (5.2) for non-analytic functions. Similar to Theorem 4.0.2, we make the following observation for mappings onto $\mathbb{R}$:

Theorem 5.0.2: *The differential $\mathrm{d}f$ of a real-valued function $f(\boldsymbol{z}) : \mathbb{A} \to \mathbb{R}$ with complex argument $\boldsymbol{z} \in \mathbb{A} \subset \mathbb{C}^n$ vanishes if and only if the vector-valued Wirtinger derivative is zero:*

$$\mathrm{d}f = 0 \Leftrightarrow \frac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{z}} = \boldsymbol{0}.$$

(5.4)

Finally, the conjugate Wirtinger derivative again points into the direction of the steepest ascent:

Corollary 5.0.1. *The steepest ascent of a real-valued function $f(\boldsymbol{z}) : \mathbb{A} \to \mathbb{R}$ with complex-valued argument $\boldsymbol{z} \in \mathbb{A} \subset \mathbb{C}^n$ is obtained for*

$$\mathrm{d}\boldsymbol{z} = \frac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{z}^*} \, \mathrm{d}s,$$

(5.5)

where $\mathrm{d}s$ is a real-valued differential.

Proof: From Theorem 5.0.1, the differential of the function f with real-valued image reads as

$$\mathrm{d}f = 2\Re\left\{\frac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{z}^{\mathrm{T}}} \, \mathrm{d}\boldsymbol{z}\right\}.$$

According to the *Cauchy-Schwarz* inequality, $\mathrm{d}\boldsymbol{z}$ has to be the conjugate of $\left(\frac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{z}^{\mathrm{T}}}\right)^{\mathrm{T}}$ times a real-valued differential and the proof is complete.

$\square$

A gradient ascent step could for example read as

$$\boldsymbol{z} \leftarrow \boldsymbol{z} + 2\frac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{z}^*} \, \mathrm{d}s$$

(5.7)

with the step-size $\mathrm{d}s$.

Examples for vector-valued Wirtinger derivatives:

- $f(\boldsymbol{z}) = \boldsymbol{z}^{\mathrm{H}} \boldsymbol{A} \boldsymbol{z}$ $\quad \dfrac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{z}} = \boldsymbol{A}^{\mathrm{T}} \boldsymbol{z}^*$

$$\dfrac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{z}^{\mathrm{T}}} = \boldsymbol{z}^{\mathrm{H}} \boldsymbol{A}$$

$$\dfrac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{z}^*} = \boldsymbol{A} \boldsymbol{z}$$

$$\dfrac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{z}^{\mathrm{H}}} = \boldsymbol{z}^{\mathrm{T}} \boldsymbol{A}^{\mathrm{T}}$$

6. Matrix-Valued Derivatives of Real-Valued Scalar-Fields

In this section, derivatives of real-valued scalar-fields are investigated. Common representatives of such scalar fields in communications and signal processing are trace or determinant expressions.

Definition 6.0.1. *Let $f : \mathbb{C}^{n \times n} \to \mathbb{R}$ denote a functional acting as a map $\boldsymbol{A} \mapsto f(\boldsymbol{A})$. Then the derivative of $f(\boldsymbol{A})$ with respect to the matrix $\boldsymbol{A}$ returns a matrix-valued function whose entry in the m-th row and n-th column reads as*

$$\left[\frac{\partial}{\partial \boldsymbol{A}} f(\boldsymbol{A}) \right]_{m,n} = \frac{\partial}{\partial [\boldsymbol{A}]_{m,n}} f(\boldsymbol{A}). \tag{6.1}$$

Sometimes the functional f is a composition of several operations where the outer one is a linear operator like the trace-operator for example. In such a case, the partial derivative operator and this outer linear operator may be interchanged as their commutator vanishes:

$$\frac{\partial}{\partial t} \operatorname{tr}(\boldsymbol{A}(t)) = \operatorname{tr}\left(\frac{\partial}{\partial t} \boldsymbol{A}(t) \right).$$

Clearly, derivatives of matrices with respect to scalars turn then out to be necessary and will therefore be discussed now.

The derivative of a matrix $\boldsymbol{A}(t)$ with respect to the variable t the matrix depends on follows from the element-wise application of the partial derivative operator onto the entries of $\boldsymbol{A}$. Hence, the element in the m-th row and n-th column of the derivative reads as

$$\left[\frac{\partial}{\partial t} \boldsymbol{A}(t) \right]_{m,n} = \frac{\partial}{\partial t} [\boldsymbol{A}(t)]_{m,n} . \tag{6.3}$$

If t is complex valued, then the partial derivative operator denotes the Wirtinger derivative. Equivalently, t may stand for an element of the matrix $\boldsymbol{A}(t)$. For example, if $t = [\boldsymbol{A}]_{m,n} = a_{m,n}$ we have

$$\frac{\partial}{\partial a_{m,n}} \boldsymbol{A} = \boldsymbol{e}_m \boldsymbol{e}_n^{\mathrm{T}}, \tag{6.4}$$

where $\boldsymbol{e}_m$ denotes the m-th canonical unit vector of appropriate dimension the elements of which are all zero except for the one in the m-th row. For the following examples, all matrices are assumed

to be constant and mutually independent. Moreover, no special structure or symmetry is assumed for them.

Examples:

- $f(\boldsymbol{A}) = \mathrm{tr}(\boldsymbol{A})$

$$\frac{\partial f(\boldsymbol{A})}{\partial \boldsymbol{A}} = \sum_{k=1}^{n} \sum_{\ell=1}^{n} \boldsymbol{e}_k \boldsymbol{e}_\ell^{\mathrm{T}} \, \mathrm{tr}\left(\frac{\partial \boldsymbol{A}}{\partial [\boldsymbol{A}]_{k,\ell}} \right) = \sum_{k=1}^{n} \sum_{\ell=1}^{n} \boldsymbol{e}_k \boldsymbol{e}_\ell^{\mathrm{T}} \, \mathrm{tr}\left(\boldsymbol{e}_k \boldsymbol{e}_\ell^{\mathrm{T}} \right)$$

$$= \sum_{k=1}^{n} \sum_{\ell=1}^{n} \boldsymbol{e}_k \boldsymbol{e}_\ell^{\mathrm{T}} \boldsymbol{e}_\ell^{\mathrm{T}} \boldsymbol{e}_k = \sum_{k=1}^{n} \boldsymbol{e}_k \boldsymbol{e}_k^{\mathrm{T}} = \mathbf{I}_n$$

- $f(\boldsymbol{A}) = \mathrm{tr}(\boldsymbol{AB})$

$$\frac{\partial f(\boldsymbol{A})}{\partial \boldsymbol{A}} = \sum_{k=1}^{n} \sum_{\ell=1}^{n} \boldsymbol{e}_k \boldsymbol{e}_\ell^{\mathrm{T}} \, \mathrm{tr}\left(\frac{\partial \boldsymbol{AB}}{\partial [\boldsymbol{A}]_{k,\ell}} \right) = \sum_{k=1}^{n} \sum_{\ell=1}^{n} \boldsymbol{e}_k \boldsymbol{e}_\ell^{\mathrm{T}} \, \mathrm{tr}\left(\boldsymbol{e}_k \boldsymbol{e}_\ell^{\mathrm{T}} \boldsymbol{B} \right)$$

$$= \sum_{k=1}^{n} \sum_{\ell=1}^{n} \boldsymbol{e}_k \boldsymbol{e}_\ell^{\mathrm{T}} \boldsymbol{e}_\ell^{\mathrm{T}} \boldsymbol{B} \boldsymbol{e}_k = \sum_{k=1}^{n} \sum_{\ell=1}^{n} \boldsymbol{e}_k \boldsymbol{e}_\ell^{\mathrm{T}} [\boldsymbol{B}]_{l,k} = \boldsymbol{B}^{\mathrm{T}}$$

Many information theoretic expressions involve the determinant-operator. For the derivative of them, the following proposition holds:

Proposition 6.0.1. *The derivative of the determinant of a matrix $\boldsymbol{A}(t)$ which depends on a parameter t with respect to this parameter reads as*

$$\frac{\partial}{\partial t} \det(\boldsymbol{A}(t)) = \det(\boldsymbol{A}) \, \mathrm{tr}\left[\boldsymbol{A}^{-1} \frac{\partial \boldsymbol{A}(t)}{\partial t} \right]. \tag{6.5}$$

Proof: We have

$$\frac{\partial}{\partial t} \det(\boldsymbol{A}(t)) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left[\det(\boldsymbol{A}(t + \Delta t)) - \det(\boldsymbol{A}(t)) \right]$$

$$= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left[\det(\boldsymbol{A}(t) + \Delta t \boldsymbol{B}(t)) - \det(\boldsymbol{A}(t)) \right] \tag{6.6}$$

$$= \det(\boldsymbol{A}(t)) \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left[\det(\mathbf{I}_n + \Delta t \boldsymbol{A}^{-1}(t) \boldsymbol{B}(t)) - 1 \right],$$

where $\boldsymbol{B}(t) = \frac{\partial}{\partial t} \boldsymbol{A}(t)$. Making use of the *Schur*-decomposition [Golub and Loan, 1991], we can rewrite $\boldsymbol{A}^{-1}(t) \boldsymbol{B}(t)$ to $\boldsymbol{A}^{-1}(t) \boldsymbol{B}(t) = \boldsymbol{Q} \boldsymbol{\Lambda} \boldsymbol{Q}^{\mathrm{H}}$ with unitary $\boldsymbol{Q}$ and the upper triangular matrix $\boldsymbol{\Lambda}$ the diagonal values of which are the eigenvalues λ_i, $i \in \{1, \ldots, n\}$. We get

$$\frac{\partial}{\partial t} \det(\boldsymbol{A}(t)) = \det(\boldsymbol{A}(t)) \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left[\prod_{i=1}^{n} (1 + \Delta t \lambda_i) - 1 \right]$$

$$= \det(\boldsymbol{A}(t)) \sum_{i=1}^{n} \lambda_i = \det(\boldsymbol{A}(t)) \, \mathrm{tr}(\boldsymbol{\Lambda}) =$$

$$= \det(\boldsymbol{A}(t)) \, \mathrm{tr}\left[\boldsymbol{A}^{-1}(t) \boldsymbol{B}(t) \right].$$

This completes the proof.

$\square$

From differentiating the identity matrix $\mathbf{I}_n = \boldsymbol{A}(t) \boldsymbol{A}^{-1}(t)$ with respect to t by means of the product rule, we obtain the following proposition:

Proposition 6.0.2. *The derivative of the inverse of a matrix $\boldsymbol{A}(t)$ with respect to the parameter t reads as*

$$\frac{\partial}{\partial t}\boldsymbol{A}^{-1}(t) = -\boldsymbol{A}^{-1}(t)\frac{\partial \boldsymbol{A}(t)}{\partial t}\boldsymbol{A}^{-1}(t). \tag{6.7}$$

Bibliography

[Golub and Loan, 1991] Golub, G. H. and Loan, C. V. (1991). *Matrix Computations*. Johns Hopkins University Press.

[Henrici, 1974] Henrici, P. (1974). *Applied and Computational Complex Analysis (Volume I)*, volume 1. Wiley-Interscience.

[Krantz, 1992] Krantz, S. G. (1992). *Function Theory of Several Complex Variables*. American Mathematical Society.

[Lang, 1993] Lang, S. (1993). *Complex Analysis (Graduate Texts in Mathematics)*. Springer.

[Palka, 1991] Palka, B. P. (1991). *An Introduction to Complex Function Theory (Undergraduate Texts in Mathematics)*. Springer.

[Schmieder, 1993] Schmieder, G. (1993). *Grundkurs Funktionentheorie*. Teubner Verlag.

[Spiegel, 1974] Spiegel, M. (1974). *THEORY AND PROBLEMS OF Complex Variables with an introduction to CONFORMAL MAPPING and its applications*. McGraw Hill Book Company.

[Trapp, 1996] Trapp, H. (1996). *Funktionentheorie einer Veränderlichen*. Universitäẗatsverlag Osnabr̈uck bei V&R unipress.

[Wirtinger, 1926] Wirtinger, W. (1926). Zur formalen Theorie der Funktionen von mehr komplexen Ver̈anderlichen.