1 Introduction

1.1 Introduction to the Book

To solve increasingly complicated open research problems, it is crucial to develop useful mathematical tools. Often, the task of a researcher or an engineer is to find the optimal values of unknown parameters that can be represented by complex-valued matrices. One powerful tool for finding the optimal values of complex-valued matrices is to calculate the derivatives with respect to these matrices. In this book, the main focus is on complex-valued matrix calculus because the theory of *real-valued* matrix derivatives has been thoroughly covered already in an excellent manner in Magnus and Neudecker (1988). The purpose of this book is to provide an introduction to the area of complex-valued matrix derivatives and to show how they can be applied as a tool for solving problems in signal processing and communications.

The framework of complex-valued matrix derivatives can be used in the optimization of systems that depend on complex design parameters in areas where the unknown parameters are complex-valued matrices with independent components, or where they belong to sets of matrices with certain structures. Many of the results discussed in this book are summarized in tabular form, so that they are easily accessible. Several examples taken from recently published material show how signal processing and communication systems can be optimized using complex-valued matrix derivatives. Note that the differentiation procedure is usually not sufficient to solve such problems completely; however, it is often an essential step toward finding the solution to the problem.

In many engineering problems, the unknown parameters are complex-valued matrices, and often, the task of the system designer is to find the values of these complex parameters, which optimize a certain scalar real-valued objective function. For solving these kinds of optimization problems, one approach is to find necessary conditions for optimality. Chapter 3 shows that when a scalar real-valued function depends on a complex-valued matrix variable, the necessary conditions for optimality can be found by setting the derivative of the function with respect to the complex-valued matrix variable or its complex conjugate to zero. It will also be shown that the direction of the maximum rate of change of a real-valued scalar function, with respect to the complex-valued matrix variable, is given by the derivative of the function with respect to the *complex conjugate* of the complex-valued input matrix variable. This result has important applications in, for example, complex-valued adaptive filters.

This book presents a comprehensive theory on how to obtain the derivatives of scalar-, vector-, and matrix-valued functions with respect to complex matrix variables. The theory of finding complex-valued matrix derivatives with respect to unpatterned matrices is based on the *complex differential* of the function of interest. The method of using differentials is substantially different from the component-wise approach.[1] A key idea when using complex differentials is to treat the *differential* of the complex and the complex conjugate variables as *independent*. This theory will be applied to derive useful matrix derivatives that can be used, for example, in signal processing and communications.

The complex Hessian matrix will be defined for complex scalar, vector, and matrix functions, and how this matrix can be obtained from the second-order differential of these functions is shown. Hessians are useful, for example, to check whether a stationary point is a saddle point, a local minimum, or a local maximum; Hessians can also be used to speed up the convergence of iterative algorithms.

A systematic theory on how to find *generalized complex-valued matrix derivatives* is presented. These are derivatives of complex-valued matrix functions with respect to matrices that belong to a set of complex-valued matrices, which might contain certain dependencies among the matrix elements. Such matrices include Hermitian, symmetric, diagonal, skew-symmetric, and skew-Hermitian. The theory of manifolds is used to find generalized complex-valued matrix derivatives. One key point of this theory is the requirement that the function, which spans all matrices within the set under consideration, is diffeomorphic; this function will be called the *parameterization function*. Several examples show how to find generalized complex-valued matrix derivatives with respect to matrices belonging to sets of matrices that are relevant for signal processing and communications.

Various applications from signal processing and communications are presented throughout the book. The last chapter is dedicated to various applications of complex-valued matrix derivatives.

1.2 Motivation for the Book

Complex signals appear in many parts of signal processing and communications. Good introductions to complex-valued signal processing can be found in Mandic and Goh (2009) and Schreier and Scharf (2010). One area where optimization problems with complex-valued matrices appear is digital communications, in which digital filters may contain complex-valued coefficients (Paulraj, Nabar, & Gore 2003). Other areas include analysis of power networks and electric circuits (González-Vázquez 1988); control theory (Alexander 1984); adaptive filters (Hanna & Mandic 2003; Diniz 2008); resource management (Han & Liu 2008); sensitivity analysis (Fränken 1997; Tsipouridou & Liavas 2008); and acoustics, optics, mechanical vibrating systems, heat con-

[1] In the author's opinion, the current approach of complex-valued matrix derivatives is preferred because it often leads to shorter and simpler calculations.

duction, fluid flow, and electrostatics (Kreyszig 1988). Convex optimization, in which the unknown parameters might be complex-valued, is treated in Boyd and Vandenberghe (2004) and Palomar and Eldar (2010). Usually, using complex-valued matrices leads to fewer computations and more compact expressions compared with treating the real and imaginary parts as two independent real-valued matrices. The complex-valued approach is general and usually easier to handle than working with the real and imaginary parts separately, because the complex matrix variable and its complex conjugate should be treated as independent variables when complex-valued matrix derivatives are calculated.

One of the main reasons why *complex-valued matrix derivatives* are so important is that necessary conditions for optimality can be found through these derivatives. By setting the complex-valued matrix derivative of the objective function equal to zero, necessary conditions for optimality are found. The theory of complex-valued matrix derivatives and the generalized complex-valued matrix derivatives are useful tools for researchers and engineers interested in designing systems in which the parameters are complex-valued matrices. The theory of *generalized* complex-valued matrix derivatives is particularly suited for problems with some type of structure within the unknown matrix of the optimization problem under consideration. Examples of such structured matrices include complex-valued diagonal, symmetric, skew-symmetric, Hermitian, skew-Hermitian, orthogonal, unitary, and positive semidefinite matrices. Finding derivatives with respect to complex-valued structured matrices is related to the field of manifolds. The theory of manifolds is a part of mathematics involving generalized derivatives on special geometric constructions spanned by so-called diffeomorphic functions (i.e., smooth invertible functions with a smooth inverse), which map the geometric construction back to a space with independent components. Optimization over such complex-valued constrained matrix sets can be done by using the theory of generalized matrix derivatives.

Complex-valued matrix derivatives are often used as a tool for solving problems in signal processing and communications. In the next section, a short overview of some of the literature on matrix derivatives is presented.

1.3 Brief Literature Summary

An early contribution to real-valued symbolic matrix calculus is found in Dwyer and Macphail (1948), which presents a basic treatment of matrix derivatives. Matrix derivatives in multivariate analysis are presented in Dwyer (1967). Another contribution is given in Nel (1980), which emphasizes the statistical applications of matrix derivatives.

The original work (Wirtinger 1927) showed that the complex variable and its complex conjugate can be treated as independent variables when finding derivatives. An introduction on how to find the Wirtinger calculus with respect to complex-valued scalars and vectors can be found in Fischer (2002, Appendix A). In Brandwood (1983), a theory is developed for finding derivatives of complex-valued scalar functions with respect to complex-valued *vectors*. It is argued in Brandwood (1983) that it is better to use the complex-valued vector and its complex conjugate as input variables instead of the real and imaginary parts of the vector – the main reason being that the complex-valued approach often leads to a simpler approach that requires fewer calculations than the method that treats the real and imaginary parts explicitly. Mandic and Goh (2009, p. 20) mention that the complex-valued representation may not always have a real physical interpretation; however, the complex framework is general and more mathematically tractable than working on the real and imaginary parts done separately.

An introduction to matrix derivatives, which focuses on component-wise derivatives, and to the Kronecker product is found in Graham (1981). Moon and Stirling (2000, Appendix E) focused on component-wise treatment of both real-valued and complex-valued matrix derivatives. Several useful results on complex-valued matrices are collected into Trees (2002, Appendix A), which also contains a few results on matrix calculus for which a component-wise treatment was used.

Magnus and Neudecker (1988) give a very solid treatment of real-valued matrices with independent components. However, they do not consider the case of formal derivatives, where the differential of the complex-valued matrix and the differential of its complex conjugate should be treated as *independent*; moreover, they do not treat the case of finding derivatives with respect to complex-valued *patterned matrices* (i.e., matrices containing certain structures). The problem of finding derivatives with respect to *real-valued* matrices containing independent elements is well known and has been studied, for example, in Harville (1997) and Minka (December 28, 2000). A substantial collection of derivatives in relation to real-valued vectors and matrices can be found in Lütkepohl (1996, Chapter 10).

Various references give brief treatments of the case of finding derivatives of real-valued scalar functions that depend on complex-valued vectors (van den Bos 1994*a*; Hayes 1996, Section 2.3.10; Haykin 2002, Appendix B; Sayed 2008, Background Material, Chapter C). A systematic and simple way to find derivatives with respect to *unpatterned* complex-valued *matrices* is presented in Hjørungnes and Gesbert (2007*a*).

Two online publications (Kreutz-Delgado 2008) and (Kreutz-Delgado 2009) give an introduction to real- and complex-valued derivatives with respect to vectors. Both first- and second-order derivatives are studied in these references. Two Internet sites with useful material on matrix derivatives are *The Matrix Cookbook* (Petersen & Pedersen 2008) and *The Matrix Reference Manual* (Brookes, July 25, 2009).

Hessians (second-order derivatives) of scalar functions of complex vectors are studied in van den Bos (1994*a*). The theory for finding Hessian matrices of scalar complex-valued function with respect to *unpatterned* complex-valued matrices and its complex conjugate is developed in Hjørungnes and Gesbert (2007*b*).

The theory for finding derivatives of real-valued functions that depend on patterned real-valued matrices is developed in Tracy and Jinadasa (1988). In Hjørungnes and Palomar (2008*b*), the theory for finding derivatives of functions that depend on complex-valued *patterned* matrices is studied; this was extended in Hjørungnes and Palomar (2008*a*), where the connections to manifolds are exploited. In Palomar and Verdú (2006), derivatives of certain scalar functions with respect to complex-valued matrices are discussed, and some results for complex-valued scalar functions with respect to matrices that contain dependent elements are presented. Vaidyanathan et al. (2010, Chapter 20), presents a treatment of real- and complex-valued matrix derivatives; however, it is based on component-wise developments. Some results on derivatives with respect to patterned matrices are presented in Vaidyanathan et al. (2010, Chapter 20).

1.4 Brief Outline

Some of the important notations used in this book and various useful formulas are discussed in Chapter 2. These items provide background material for later chapters. A classification of complex variables and functions is also presented in Chapter 2, which includes a discussion of the differences between analytic functions – subject matter usually studied in mathematical courses for engineers, and non-analytic functions, which are encountered when dealing with practical engineering problems of complex variables.

In Chapter 3, the complex differential is introduced. Based on the complex differential, the definition of the derivatives of complex-valued matrix functions with respect to the unpatterned complex-valued matrix variable and its complex conjugate is introduced. In addition, a procedure showing how the derivatives can be found from the differential of a function when the complex matrix variable contains independent elements is presented in Chapter 3. This chapter also contains several important results stated in theorems, such as the chain rule and necessary conditions for optimality for real-valued scalar functions.

Chapter 4 states several results in tables and shows how most of these results can be derived for nine different types of functions. These nine function types result when the input and the output of the function take the form of a scalar, a vector, or a matrix.

The Hessian matrix of complex-valued scalar, vector, and matrix functions dependent on complex matrices is defined in Chapter 5, which shows how this Hessian matrix can be obtained from the second-order differential. Hessian matrices can, for example, be used to speed up convergence of iterative algorithms, to study the convexity and concavity of an objective function, and to perform stability analysis of iterative algorithms.

Often, in signal processing and communications, the challenge is to find a matrix that optimizes a problem when the matrix is constrained to belong to a certain set, such as Hermitian matrices or symmetric matrices. For solving such types of problems, derivatives associated with matrices belonging to these sets are useful. These types of derivatives are called *generalized complex-valued matrix derivatives*, and a theory for finding such derivatives is presented in Chapter 6.

In Chapter 7, various applications taken from signal processing and communications are presented to show how complex-valued matrix derivatives can be used as a tool to solve research problems in these two fields.

After the seven chapters, references and the index follow.

2 Background Material

2.1 Introduction

In this chapter, most of the notation used in this book will be introduced. It is *not* assumed that the reader is familiar with topics such as Kronecker product, Hadamard product, or vectorization operator. Therefore, this chapter defines these concepts and gives some of their properties. The current chapter also provides background material for matrix manipulations that will be used later in the book. However, it contains just the minimum of material that will be used later because many excellent books in linear algebra are available for the reader to consult (Gantmacher 1959*a*–1959*b*; Horn & Johnson 1985; Strang 1988; Magnus & Neudecker 1988; Golub & van Loan 1989; Horn & Johnson 1991; Lütkepohl 1996; Harville 1997; Bernstein 2005).

This chapter is organized as follows: Section 2.2 introduces the basic notation and classification used for complex-valued variables and functions. A discussion of the differences between analytic and non-analytic functions is presented in Section 2.3. Basic matrix-related definitions are provided in Section 2.4. Several results involving matrix manipulations used in later chapters are found in Section 2.5. Section 2.6 offers exercises related to the material included in this chapter. Theoretical derivations and computer programming in `MATLAB` are topics of these exercises.

2.2 Notation and Classification of Complex Variables and Functions

Denote $\mathbb{R}$ and $\mathbb{C}$ the sets of the real and complex numbers, respectively, and define $\mathbb{Z}_N \triangleq \{0, 1, \ldots, N-1\}$. The notation used for the two matrices consisting entirely of zeros and ones is $\mathbf{0}_{N\times Q}$ and $\mathbf{1}_{N\times Q}$, respectively, where the size of the matrices is indicated by the subindex to be $N \times Q$.

The following conventions are always used in this book:

- Scalar quantities are denoted by lowercase symbols.
- Vector quantities are denoted by lowercase boldface symbols.
- Matrix quantities are denoted by capital boldface symbols.

Table 2.1 Symbols and sizes of the most frequently used variables and functions.

Symbol	z	$\boldsymbol{z}$	$\boldsymbol{Z}$	f	$\boldsymbol{f}$	$\boldsymbol{F}$
Size	1×1	$N \times 1$	$N \times Q$	1×1	$M \times 1$	$M \times P$

2.2.1 Complex-Valued Variables

A function's complex input argument can be a scalar, denoted z, a vector, denoted $\boldsymbol{z}$, or a matrix, denoted $\boldsymbol{Z}$.

Let the symbol z denote a complex scalar variable, and let the real and imaginary part of z be denoted by x and y, respectively, then

$$z = x + \jmath y, \tag{2.1}$$

where $\jmath$ is the imaginary unit, and $\jmath^2 = -1$. The absolute value of the complex number z is denoted by $|z|$.

The real and imaginary operators return the real and imaginary parts of the input matrix, respectively. These operators are denoted by $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$. If $\boldsymbol{Z} \in \mathbb{C}^{N\times Q}$ is a complex-valued matrix, then

$$\boldsymbol{Z} = \mathrm{Re}\{\boldsymbol{Z}\} + \jmath\, \mathrm{Im}\{\boldsymbol{Z}\}, \tag{2.2}$$

$$\boldsymbol{Z}^* = \mathrm{Re}\{\boldsymbol{Z}\} - \jmath\, \mathrm{Im}\{\boldsymbol{Z}\}, \tag{2.3}$$

where $\mathrm{Re}\{\boldsymbol{Z}\} \in \mathbb{R}^{N\times Q}$, $\mathrm{Im}\{\boldsymbol{Z}\} \in \mathbb{R}^{N\times Q}$, and the operator $(\cdot)^*$ denote the complex conjugate of the matrix it is applied to. The real and imaginary operators can be expressed as

$$\mathrm{Re}\{\boldsymbol{Z}\} = \frac{1}{2}(\boldsymbol{Z} + \boldsymbol{Z}^*), \tag{2.4}$$

$$\mathrm{Im}\{\boldsymbol{Z}\} = \frac{1}{2\jmath}(\boldsymbol{Z} - \boldsymbol{Z}^*). \tag{2.5}$$

2.2.2 Complex-Valued Functions

For complex-valued functions, the following conventions are always used in this book:

- If the function returns a scalar, then a lowercase symbol is used, for example, f.
- If the function returns a vector, then a lowercase boldface symbol is used, for example, $\boldsymbol{f}$.
- If the function returns a matrix, then a capital boldface symbol is used, for example, $\boldsymbol{F}$.

Table 2.1 shows the sizes and symbols of the variables and functions most frequently used in the part of the book that treats complex matrix derivatives with independent components. Note that $\boldsymbol{F}$ covers all situations because scalars f and vectors $\boldsymbol{f}$ are special cases of a matrix. In the sequel, however, the three types of functions are distinguished as scalar, vector, or matrix because, as we shall see in Chapter 4, different definitions of the derivatives, based on type of functions, are found in the literature.

2.3 Analytic versus Non-Analytic Functions

Let the symbol $\subseteq$ mean subset of, and $\subset$ proper subset of. Mathematical courses on complex functions for *engineers* often involve only the analysis of analytic functions (Kreyszig 1988, p. 738) defined as follows:

Definition 2.1 (Analytic Function) *Let $D \subseteq \mathbb{C}$ be the domain*[1] *of definition of the function $f : D \rightarrow \mathbb{C}$. The function f is an* analytic function in the domain *D if $\lim_{\Delta z \rightarrow 0} \frac{f(z + \Delta z) - f(z)}{\Delta z}$ exists for all $z \in D$.*

If $f(z)$ satisfies the Cauchy-Riemann equations (Kreyszig 1988, pp. 740–743), then it is analytic. A function that is analytic is also named complex differentiable, holomorphic, or regular. The Cauchy-Riemann equations for the scalar function f can be formulated as a single equation in the following way:

$$\frac{\partial}{\partial z^*} f = 0. \tag{2.6}$$

From (2.6), it is seen that any analytic function f is *not* dependent on the variable z^*. This can also be seen from Theorem 1 in Kreyszig (1988, p. 804), which states that any analytic function $f(z)$ can be written as a power series[2] with non-negative exponents of the complex variable z, and this power series is called the Taylor series. This series does *not* contain any terms that depend on z^*. The derivative of a complex-valued scalar function in mathematical courses of complex analysis for *engineers* is often defined only for analytic functions. However, in engineering problems, the functions of interest often are *not* analytic because they are often real-valued functions. If a function is dependent only on z, as are analytic functions, and is not implicitly or explicitly dependent on z^*, then this function *cannot* in general be real-valued; a function can be real-valued only if the imaginary part of f vanishes, and this is possible only if the function also depends on terms that depend on z^*. An alternative treatment for how to find the derivative of real functions dependent on complex variables other than the one used for analytic function is needed. In this book, a theory that solves this problem is provided for scalar, vector, or matrix functions and variables.

[1] If $f : \mathcal{A} \rightarrow \mathcal{B}$, then the set $\mathcal{A}$ is called the *domain* of f, the set $\mathcal{B}$ is called the *range* of f, and the set $\{f(x) \mid x \in \mathcal{A}\}$ is called the *image set* of f (Munkres 2000, p. 16).

[2] A power series in the variable $z \in \mathbb{C}$ is an infinite sum of the form $\sum_{n=0}^{\infty} a_n (z - z_0)^n$, where a_n, $z_0 \in \mathbb{C}$ (Kreyszig 1988, p. 812).

In engineering problems, the squared Euclidean distance is often used. Let $f : \mathbb{C} \to \mathbb{C}$ be defined as

$$f(z) = |z|^2 = zz^*. \tag{2.7}$$

If the traditional definition of the derivative given in Definition 2.1 is used, then the function f is not differentiable because

$$\begin{aligned} \lim_{\Delta z \to 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z} &= \lim_{\Delta z \to 0} \frac{|z_0 + \Delta z|^2 - |z_0|^2}{\Delta z} \\ &= \lim_{\Delta z \to 0} \frac{(z_0 + \Delta z)(z_0^* + (\Delta z)^*) - z_0 z_0^*}{\Delta z} \\ &= \lim_{\Delta z \to 0} \frac{(\Delta z) z_0^* + z_0 (\Delta z)^* + \Delta z (\Delta z)^*}{\Delta z}, \end{aligned} \tag{2.8}$$

and this limit does *not exist*, because different values are found depending on how Δz is approaching 0. Let $\Delta z = \Delta x + \jmath \Delta y$. First, let Δz approach 0 such that $\Delta x = 0$, then the last fraction in (2.8) is

$$\frac{\jmath (\Delta y) z_0^* - \jmath z_0 \Delta y + (\Delta y)^2}{\jmath \Delta y} = z_0^* - z_0 - \jmath \Delta y, \tag{2.9}$$

which approaches $z_0^* - z_0 = -2\jmath \operatorname{Im}\{z_0\}$, when $\Delta y \to 0$. Second, let Δz approach 0 such that $\Delta y = 0$, then the last fraction in (2.8) is

$$\frac{(\Delta x) z_0^* + z_0 \Delta x + (\Delta x)^2}{\Delta x} = z_0 + z_0^* + \Delta x, \tag{2.10}$$

which approaches $z_0 + z_0^* = 2 \operatorname{Re}\{z_0\}$ when $\Delta x \to 0$. For an arbitrary complex number z_0, in general, $2 \operatorname{Re}\{z_0\} \neq -2\jmath \operatorname{Im}\{z_0\}$. This means that the function $f(z) = |z|^2 = zz^*$ is *not* differentiable when the commonly encountered definition given in Definition 2.1 is used, and, hence, f is not analytic.

Two alternative ways (Hayes 1996, Subsection 2.3.10) are known for finding the derivative of a scalar real-valued function $f \in \mathbb{R}$ with respect to the unknown complex-valued matrix variable $\boldsymbol{Z} \in \mathbb{C}^{N \times Q}$. The first way is to rewrite f as a function of the real $\boldsymbol{X}$ and imaginary parts $\boldsymbol{Y}$ of the complex variable $\boldsymbol{Z}$, and then to find the derivatives of the rewritten function with respect to these two independent real variables, $\boldsymbol{X}$ and $\boldsymbol{Y}$, separately. Notice that NQ independent complex unknown variables in $\boldsymbol{Z}$ correspond to $2NQ$ independent real variables in $\boldsymbol{X}$ and $\boldsymbol{Y}$. The second way to deal with this problem, which is more elegant and is used in this book, is to treat the differentials of the variables $\boldsymbol{Z}$ and $\boldsymbol{Z}^*$ as independent, in the way that will be shown by Lemma 3.1. Chapter 3 shows that the derivative of f with respect to $\boldsymbol{Z}$ and $\boldsymbol{Z}^*$ can be identified by the differential of f.

Complex numbers cannot be ordered as real numbers can. Therefore, the objective functions of interest, when dealing with engineering problems, are usually real valued in such a way that it makes sense to minimize or maximize them. If a real-valued function depends on a complex matrix $\boldsymbol{Z}$, it must also be explicitly or implicitly dependent on $\boldsymbol{Z}^*$, such that the result is real (see also the discussion following (2.6)). A real-valued

Table 2.2 Classification of functions.

Function type	$z, z^* \in \mathbb{C}$	$\boldsymbol{z}, \boldsymbol{z}^* \in \mathbb{C}^{N\times 1}$	$\boldsymbol{Z}, \boldsymbol{Z}^* \in \mathbb{C}^{N\times Q}$
Scalar function $f \in \mathbb{C}$	$f(z, z^*)$ $f : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$	$f(\boldsymbol{z}, \boldsymbol{z}^*)$ $f : \mathbb{C}^{N\times 1} \times \mathbb{C}^{N\times 1} \to \mathbb{C}$	$f(\boldsymbol{Z}, \boldsymbol{Z}^*)$ $f : \mathbb{C}^{N\times Q} \times \mathbb{C}^{N\times Q} \to \mathbb{C}$
Vector function $\boldsymbol{f} \in \mathbb{C}^{M\times 1}$	$\boldsymbol{f}(z, z^*)$ $\boldsymbol{f} : \mathbb{C} \times \mathbb{C} \to \mathbb{C}^{M\times 1}$	$\boldsymbol{f}(\boldsymbol{z}, \boldsymbol{z}^*)$ $\boldsymbol{f} : \mathbb{C}^{N\times 1} \times \mathbb{C}^{N\times 1} \to \mathbb{C}^{M\times 1}$	$\boldsymbol{f}(\boldsymbol{Z}, \boldsymbol{Z}^*)$ $\boldsymbol{f} : \mathbb{C}^{N\times Q} \times \mathbb{C}^{N\times Q} \to \mathbb{C}^{M\times 1}$
Matrix function $\boldsymbol{F} \in \mathbb{C}^{M\times P}$	$\boldsymbol{F}(z, z^*)$ $\boldsymbol{F} : \mathbb{C} \times \mathbb{C} \to \mathbb{C}^{M\times P}$	$\boldsymbol{F}(\boldsymbol{z}, \boldsymbol{z}^*)$ $\boldsymbol{F} : \mathbb{C}^{N\times 1} \times \mathbb{C}^{N\times 1} \to \mathbb{C}^{M\times P}$	$\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*)$ $\boldsymbol{F} : \mathbb{C}^{N\times Q} \times \mathbb{C}^{N\times Q} \to \mathbb{C}^{M\times P}$

Adapted from Hjørungnes and Gesbert (2007*a*). © 2007 IEEE.

function can consist of several terms; it is possible that some of these terms are complex valued, even though their sum is real.

The main types of functions used throughout this book, when working with complex-valued matrix derivatives with independent components, can be classified as in Table 2.2. The table shows that all functions depend on a complex variable and the complex conjugate of the same variable, and the reason for this is that the complex *differential* of the variables $\boldsymbol{Z}$ and $\boldsymbol{Z}^*$ should be treated independently. When the function has two complex input variables of the same size (e.g., $\boldsymbol{F} : \mathbb{C}^{N\times Q} \times \mathbb{C}^{N\times Q} \to \mathbb{C}^{M\times P}$ for the general case), then two input variables should be the complex conjugate of each other. This means that they cannot be chosen independently of each other. However, in Lemmas 3.1 and 6.1, it will be shown that the *differentials* of the two input matrix variables $\boldsymbol{Z}$ and $\boldsymbol{Z}^*$ are independent. The convention of using both a complex variable and its complex conjugate explicitly in the function definition was used in Brandwood (1983). When evaluating, for example, the most general function in Table 2.2 (i.e., $\boldsymbol{F} : \mathbb{C}^{N\times Q} \times \mathbb{C}^{N\times Q} \to \mathbb{C}^{M\times P}$), the notation adapted is that the two complex-valued input variables should be the complex conjugates of each other. Hence, the two input arguments of $\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*)$ are a function of each other, but as will be seen in Lemma 3.1, the *differentials* of the two input variables $\boldsymbol{Z}$ and $\boldsymbol{Z}^*$ are independent. When working with complex-valued matrix derivatives in later chapters, we will see that complex differentials are very important.

Definition 2.2 (Formal Derivatives) *Let $z = x + \jmath y$, where $x, y \in \mathbb{R}$, then the* formal derivatives*, with respect to z and z^* of $f(z_0)$ at $z_0 \in \mathbb{C}$ or* Wirtinger derivatives (Wirtinger 1927)*, are defined as*

$$\frac{\partial}{\partial z} f(z_0) = \frac{1}{2}\left(\frac{\partial}{\partial x} f(z_0) - \jmath \frac{\partial}{\partial y} f(z_0)\right), \tag{2.11}$$

$$\frac{\partial}{\partial z^*} f(z_0) = \frac{1}{2}\left(\frac{\partial}{\partial x} f(z_0) + \jmath \frac{\partial}{\partial y} f(z_0)\right). \tag{2.12}$$

When finding $\frac{\partial}{\partial z} f(z_0)$ and $\frac{\partial}{\partial z^*} f(z_0)$, the variables z and z^* are treated as independent variables (Brandwood 1983, Theorem 1).