Çağatay Candan

Properly Handling Complex Differentiation in Optimization and Approximation Problems

Functions of complex variables arise frequently in the formulation of signal processing problems. The basic calculus rules on differentiation and integration for functions of complex variables resemble, but are not identical to, the rules of their real variable counterparts. On the contrary, the standard calculus rules on differentiation, integration, series expansion, and so on are the special cases of the complex analysis with the restriction of the complex variable to the real line. The goal of this lecture note is to review the fundamentals of the functions of complex variables, highlight the differences and similarities with their real variable counterparts, and study the complex differentiation operation with the optimization and approximation applications in mind. More specifically, the take-home result of this lecture note is to understand the differentiation with respect to the conjugate variable $(\partial/\partial\bar{z})f(z,\bar{z})$, which is known as *Wirtinger calculus*, and its application in optimization and approximation problems.

Relevance

Complex analysis is a rich and interesting topic with close ties to the foundation of our profession. To illustrate its foundational nature, it is important to remember that the independent variable of a transfer function, e.g., $H(s)$ or $H(z)$, is a complex variable. Hence, the transfer function, a concept at the core of many signal processing operations, is nothing but a mapping via the function of a complex variable. In the early days of signal processing, in the late 1970s to early 1980s, an electrical engineering curriculum could not be considered "good" without a mandatory course on complex analysis where differentiable functions, Taylor/Laurent series, contour integration, and residue calculation, were presented in some depth to an undergraduate audience. With the progress of time, the attention of undergraduate students has been directed elsewhere, and topics of complex calculus have been eliminated from most undergraduate curricula. Today, with the exception of a very few students, even top graduate students do not know why the standard partial-fraction expansion method is called the *residue method*. Additionally, they are unable to explain the equality of $\int_{-\infty}^{\infty} \sin(\pi x)/(\pi x)\,dx = 1$, i.e., the area under sinc function, without reverting to the infamous duality property of the Fourier transform. Furthermore, they cannot discriminate between the removable singularity of $\sin(\pi z)/(\pi z)$ at $z = 0$ and the pole of $1/z$ at $z = 0$. This article is prepared not to present a remedy to all mentioned issues but to provide an overview of the most basic rules for complex-valued optimization/approximation by differentiation, which is essential for many signal processing applications. We also provide several pointers for the readers that wish further explore the topics of complex analysis on their own.

Prerequisites

The only prerequisites are the working knowledge of freshman calculus, basic signal processing theory, and some exposure to the signal processing applications to develop some feeling on the application range of the presented ideas.

Problem statement and solution

Problem statement

We consider functions in the form $f(z)$ where $z = x + jy$ is a complex-valued scalar and $j = \sqrt{-1}$ is the imaginary unit, x and y are real scalars, which are called the *real* and *imaginary parts of the complex number* z, respectively. With the well-known definitions for the complex multiplication and addition, the field of complex numbers $\mathbb{C}$ is constructed. The mapping from $z \in \mathbb{C}$ to $w = f(z) \in \mathbb{C}$ is verbally described as a complex-valued function of a complex variable.

Freshman calculus allows us to determine the optima of many problems by differentiation. Students are introduced to calculus as early as possible because of its importance and wide range of application in nearly all fields of science and engineering. Yet, the functions of complex variables, differentiation operation, and utilization of differentiation in the optimization and approximation problems with complex variables are not studied in a typical engineering curriculum. In some cases, the extension from standard calculus to complex calculus can be straightforward; however, in others, the process can be seen as tricky to an inexperienced eye.

Let $f(z) \triangleq |z|^2 = x^2 + y^2$. Since function $f(z)$ is a quadratic function of real variables x and y, one may be tempted to say that $f(z)$ is a differentiable function of z considering its quadratic nature. This assumption would indeed be correct if the function $x^2 + y^2$ is considered a function of two independent "real" variables x and y; however, it would not correct for a complex variable, z. Yet, functions known as *analytic functions*, such as $f(x) = \sin(x)$ or $f(x) = x^2$, essentially carry all of the properties (e.g., differentiation, integration, and continuity) studied for their real-valued counterparts with the replacement of x with z, along with the replacement of real arithmetic with complex arithmetic.

Digital Object Identifier 10.1109/MSP.2018.2876761
Date of publication: 26 February 2019

The main problem studied here is the optimization and approximation of the functions of complex variables. We show that differentiation with respect to $\bar{z}$, the complex conjugate of z, i.e., $\partial/\partial\bar{z}\, f(z, \bar{z})$, instead of real and imaginary parts of $z = x + jy$, simplifies writing equations, their interpretation, and reduces the dimension of the problem. To clarify the meaning of $\partial/\partial\bar{z}\, f(z, \bar{z})$, we first need to establish some understanding of the complex differentiation operation.

Solution

The function $f(z)$ can be expressed as

$$f(z)\big|_{z=x+jy} = u(x, y) + jv(x, y), \quad (1)$$

where $u(x, y)$ and $v(x, y)$ are ordinary functions of two real variables. Hence, by juxtaposing arbitrary functions of two real variables with the imaginary unit j, e.g., $u(x, y) \triangleq xy^2$, $v(x, y) \triangleq x^2/y^3$, we can establish a mapping from $z \in \mathbb{C}$ to $f(z) \in \mathbb{C}$ in the form $f(z) = u(x, y) + jv(x, y)$.

The conjugate of the variable $z = x + jy$ is defined as $\bar{z} = x - jy$. With this definition, the real and imaginary parts of z can be expressed as

$$x = \frac{z + \bar{z}}{2}, \quad y = \frac{z - \bar{z}}{2j}. \quad (2)$$

It is possible interpret the transformation from (x, y) to $(z, \bar{z})$ and vice versa, given by (2), as a change of variables for two independent variables. Hence, the mapping $f(x + jy) = u(x, y) + jv(x, y)$ can be written with the alternate variables of $z, \bar{z}$ as

$$f(z, \bar{z}) = u\Bigg(\underbrace{\frac{z + \bar{z}}{2}}_{x}, \underbrace{\frac{z - \bar{z}}{2j}}_{y}\Bigg) + jv\Bigg(\underbrace{\frac{z + \bar{z}}{2}}_{x}, \underbrace{\frac{z - \bar{z}}{2j}}_{y}\Bigg). \quad (3)$$

Careful readers must have noticed the notational abuse of using $f(z) = f(x + jy)$ and $f(z, \bar{z})$ for the same mapping. Note that the function $f(z)$ is a function of two real variables that are the real and imaginary parts of z, but the notation of $f(z)$ hides this dependence. The dependence on two variables is explicit in the equivalent definition of $f(z, \bar{z})$. We ask the reader to tolerate this minor abuse of notation, which is basically required to simplify the presentation.

Functions differentiable at $z = z_0$

A complex-valued function is said to be differentiable at $z = z_0$ if the complex-valued limit operation, shown in (4) at the bottom of the page, exists.

Here, $\Delta_z = \Delta_x + j\Delta_y$ and Δ_z approaches zero, i.e., $\Delta_z \to 0$ should be interpreted as $|\Delta_z| = \sqrt{\Delta_x^2 + \Delta_y^2} \to 0$. The limit exists if the limit value is independent of the direction where Δ_z approaches zero. For example, if Δ_z approaches zero on the real axis, i.e., $\Delta_x \to 0$ and $\Delta_y = 0$, we have

$$f'(z_0) \overset{\Delta_x \to 0,\, \Delta_y = 0}{=} \frac{\partial}{\partial x} u(x_0, y_0) + j\frac{\partial}{\partial x} v(x_0, y_0). \quad (5)$$

Similarly, if Δ_z approaches zero on the imaginary axis, i.e., $\Delta_x = 0$ and $\Delta_y \to 0$, we have

$$f'(z_0) \overset{\Delta_x = 0,\, \Delta_y \to 0}{=} \frac{\partial}{\partial y} v(x_0, y_0) - j\frac{\partial}{\partial y} u(x_0, y_0). \quad (6)$$

For the existence of the limit in (4), the right-hand sides of (5) and (6) should be identical. Equating (5) and (6), we get the Cauchy–Riemann conditions for the differentiability at $z = z_0$

$$\frac{\partial}{\partial x} u(x_0, y_0) = \frac{\partial}{\partial y} v(x_0, y_0), \quad \frac{\partial}{\partial x} v(x_0, y_0) = -\frac{\partial}{\partial y} u(x_0, y_0). \quad (7)$$

$$f'(z_0) = \lim_{\Delta_z \to 0} \frac{f(z_0 + \Delta_z) - f(z_0)}{\Delta_z} = \lim_{\Delta_z \to 0} \frac{(u(x_0 + \Delta_x, y_0 + \Delta_y) + jv(x_0 + \Delta_x, y_0 + \Delta_y)) - (u(x_0, y_0) + jv(x_0, y_0))}{\Delta_x + j\Delta_y}. \quad (4)$$

As an exercise, we suggest that readers evaluate the limit for an arbitrary approach direction of $[\alpha_x \alpha_y]^T$ by taking $\Delta_z = (\alpha_x + j\alpha_y)h$ and evaluating the limit in (4) as $h \to 0$. If the Cauchy–Riemann conditions hold, i.e., the special cases of the general approach direction for $\alpha_x = 0$ or $\alpha_y = 0$, the limit value is the same for all of the approach directions. Hence, the Cauchy–Riemann conditions are not only necessary but also sufficient for the existence of the limit defining the complex differentiation operation.

With the alternate representation of the function $f(z)$ in the form $f(z, \bar{z})$, instead of $f(x + jy)$, the partial derivatives with respect to z and $\bar{z}$ can be formally written as

$$\begin{aligned}\frac{\partial}{\partial z} f(z, \bar{z}) &= \frac{\partial f(z)}{\partial x}\frac{\partial x}{\partial z} + \frac{\partial f(z)}{\partial y}\frac{\partial y}{\partial z} \\ &= \frac{1}{2}\left(\frac{\partial f(z)}{\partial x} - j\frac{\partial f(z)}{\partial y}\right), \\ \frac{\partial}{\partial \bar{z}} f(z, \bar{z}) &= \frac{\partial f(z)}{\partial x}\frac{\partial x}{\partial \bar{z}} + \frac{\partial f(z)}{\partial y}\frac{\partial y}{\partial \bar{z}} \\ &= \frac{1}{2}\left(\frac{\partial f(z)}{\partial x} + j\frac{\partial f(z)}{\partial y}\right), \end{aligned} \quad (8)$$

where the coordinate change relations in (2) are utilized for the partial derivatives of x and y with respect to z and $\bar{z}$. The expressions in (8) are first given by Wirtinger [1], and the calculus with z and $\bar{z}$ is called the *Wirtinger calculus* in some texts [2]. This concludes the definition for $\partial/\partial\bar{z}\, f(z, \bar{z})$, a result of pivotal importance in the simplification of approximation and optimization problems with complex variables.

As a first application of $\partial/\partial\bar{z}\, f(z, \bar{z})$, we take a second look at the Cauchy–Riemann equations. For the compactness of notation, we denote the partial derivative of $f(z) = u(x, y) + jv(x, y)$ with respect to x and y as $f_x = \partial f/\partial x = u_x + jv_x$ and $f_y = \partial f/\partial y = u_y + jv_y$, respectively. With the substitution of f_x and f_y into (8), we get

$$\frac{\partial}{\partial \bar{z}} f(z, \bar{z}) = \frac{1}{2}\{u_x - v_y + j(v_x + u_y)\}. \quad (9)$$

Please note that, for the functions satisfying the Cauchy–Riemann conditions, as given in (7), the right-hand side of (9) reduces to zero. Hence, differentiability at a point requires the formal

derivative of $f(z, \bar{z})$ with respect to $\bar{z}$ vanish at that point. This simpler-to-state-and-remember condition is equivalent to the Cauchy–Riemann conditions. To highlight the differences with the real-valued calculus, we note that differentiation with respect to the conjugate variable has no counterpart in standard calculus.

Differentiability in a region requires $f(z, \bar{z})$ to be independent of $\bar{z}$ in that region. Stated differently, $f(z, \bar{z})$ should be solely a function of z, i.e., should not exhibit any dependence on $\bar{z}$ for differentiability in that region. For example, $f(z) = z^2$ is differentiable throughout the complex plane since $f(z) = z^2$ is independent of $\bar{z}$ throughout the complex plane. The function $f(z) = |z|^2 = z\bar{z}$ with the partial derivative $\partial / \partial \bar{z}\, f(z, \bar{z}) = z$ is only differentiable when $z = 0$. Yet, the same function, $|z|^2 = x^2 + y^2$, is a differentiable function of real variables x and y in the standard calculus sense. We suggest that readers keep this example in mind during the discussions of differentiation with complex/real variables.

As it is evident from the aforementioned $f(z) = |z|^2$ discussion, the complex differentiation operation is much more restrictive than differentiation for real variables. This is, in essence, due to the complex multiplication definition that results in an unavoidable interaction between the real and imaginary parts of a complex-valued function. Additionally, in stark contrast to the real-valued functions, if $f(z)$ is differentiable once in a neighborhood of $z = z_0$, the function is infinitely differentiable in the same neighborhood [3, Sec. 52]. Hence, unlike standard calculus results, the Cauchy–Riemann conditions are not only the gate to the first derivative, but also to the second and all higher-order derivatives.

Analytic functions

A function of a complex variable z is said to be an analytic function in a domain if the function is differentiable for all points of that domain. Differentiability and analyticity should be considered pointwise and neighborhood properties of a function, respectively. Hence, rather confusingly, $f(z) = |z|^2$ is differentiable only at $z = 0$, but is not analytic at any point. Analyticity of functions is especially important for the contour integration (path integration) in the complex plane [3, Sec. 50]. We do not delve into integration at all in these notes, yet we bring the analyticity definition to the attention of readers to underscore that differentiability and analyticity are not synonymous for functions of complex variables.

In the next section, we examine the application of complex differentiation in the optimization and approximation problems with some illustrative examples. We examine two classes, i.e., real-valued functions of complex variables and complex-valued functions of complex variables.

Application examples

Case 1: Real-valued functions of complex variables

This case studies the functions in the form $f(z) = u(x, y)$, i.e., the functions $f(z)|_{z=x+jy} = u(x, y) + jv(x, y)$ with a vanishing imaginary part. Such functions are called *real-valued functions of complex variables*. It should be clear that these functions do not satisfy the Cauchy–Riemann conditions, unless $f(z)$ is a constant function, i.e., $f(z) = K$. From an applications viewpoint, the cost functions associated with the optimization problems are in this form. An example is the least-squares problem with $J(\mathbf{z}) = \|\mathbf{Az} - \mathbf{b}\|^2$, where $\mathbf{A}$ is an $N \times K$ matrix ($N > K$) with complex-valued entries and vectors $\mathbf{b} \in \mathbb{C}^N$ and $\mathbf{z} \in \mathbb{C}^K$. The function $J(\mathbf{z})$ can be considered to be mapping from $\mathbb{C}^K$ to $\mathbb{R}$.

A legitimate, but inefficient way to solve such optimization problems is to reduce the problem to the optimization of real variables. For example, $J(\mathbf{z}) = \|\mathbf{Az} - \mathbf{b}\|^2$ can be written as shown in (10) at the bottom of the page.

Here, $\|\mathbf{x}\|_R^2 = \mathbf{x}^T\mathbf{x}$ and $\|\mathbf{x}\|^2 = \mathbf{x}^H\mathbf{x}$ denote the vector norms for real- and complex-valued vectors, respectively. With the introduced definitions, it is easy to see that $J(\mathbf{z}) = J(\mathrm{Re}(\mathbf{z}), \mathrm{Im}(\mathbf{z}))$, allowing for some notational abuse.

In principle, the optimization problem given by (10) can be solved without any use of complex-valued operations at the expense of doubling the dimension of the problem, i.e., by solving $2N$ equations of $2K$ real unknowns instead of N equations with K complex unknowns. We illustrate the optimization of $f(\mathbf{z})$ by using complex differentiation rules—which reduces the amount of equation writing and solving—with two examples. The first one is a bare-bones application of presented ideas given to illustrate the essentials of the procedure. The second one shows the application of the mentioned ideas in a filter design problem.

Example 1: A simple optimization problem

The problem is to find the point on the circle shown in Figure 1 such that the sum of its coordinates is minimum. As an application motivation, in many sparse signal reconstruction applications, ℓ_1 norm of the unknown vector is minimized under a constraint, which is typically an underdetermined linear equation system. Our problem is somewhat similar to these problems yet differs in the constraint. The circle in Figure 1 is centered at (2, 3) and has a radius of $\sqrt{2}$. The problem can be expressed as the minimization of the cost function $J(x, y) = x + y$ under the constraint $(x-2)^2 + (y-3)^2 = 2$.

We can use the Lagrange multiplier approach to convert the constrained optimization problem to an unconstrained one

$$L(x, y, \lambda) = x + y + \lambda((x-2)^2 + (y-3)^2 - 2), \tag{11}$$

where $L(x, y, \lambda)$ is the Lagrangian function and λ is the unknown Lagrange multiplier. By taking derivatives with respect to x, y, and λ and equating them to 0, we get the following equation system:

$$\frac{\partial}{\partial x} L(x, y, \lambda) = 1 + 2\lambda(x-2) = 0, \tag{12a}$$

$$\frac{\partial}{\partial y} L(x, y, \lambda) = 1 + 2\lambda(y-3) = 0, \tag{12b}$$

$$J(\mathrm{Re}(\mathbf{z}), \mathrm{Im}(\mathbf{z})) = \left\| \begin{bmatrix} \mathrm{Re}(\mathbf{A}) & -\mathrm{Im}(\mathbf{A}) \\ \mathrm{Im}(\mathbf{A}) & \mathrm{Re}(\mathbf{A}) \end{bmatrix} \begin{bmatrix} \mathrm{Re}(\mathbf{z}) \\ \mathrm{Im}(\mathbf{z}) \end{bmatrix} - \begin{bmatrix} \mathrm{Re}(\mathbf{b}) \\ \mathrm{Im}(\mathbf{b}) \end{bmatrix} \right\|_R^2. \tag{10}$$

$$\frac{\partial}{\partial\lambda}L(x,y,\lambda) = (x-2)^2+(y-3)^2-2=0. \quad (12c)$$

Solving for x and y from (12a) and (12b), we get $x=2-1/(2\lambda)$ and $y=3-1/(2\lambda)$. Inserting these expressions into (12c), we get $\lambda=\{-0.5,0.5\}$. Hence, we find the possible solutions as (1, 2) and (3, 4). Among these extrema, the cost is minimized by the point (1, 2), which is indicated with the green diamond marker in Figure 1. The value of the minimum cost is $J(1,2)=3$.

Figure 1 also shows constant cost curves (level curves), $J(x,y)=c$, for the cost function $J(x,y)=x+y$. It is seen in Figure 1, that as the value of $J(x,y)$ increases, the level curve "approaches" the circle. From this observation, it should be clear that the cost value for which the level curve is tangent to the circle is an extremum of the problem. Hence, the optimization problem can be solved geometrically by finding the points on the circle whose tangent line has the slope of −1. Of course, the solution by this geometric approach coincides with our earlier findings, i.e., the points of (1, 2) and (3, 4).

Now, we tackle the same problem by transforming it to the complex domain. First, we introduce $z=x+jy$ to reduce the problem to the one with a single, complex unknown. By substituting $x=\mathrm{Re}(z)=((z+\bar{z})/2)$ and $y=\mathrm{Im}(z)=((z-\bar{z})/j2)$ into (11), we get

$$\begin{aligned}L(z,\bar{z},\lambda) &= \frac{z+\bar{z}}{2}+\frac{z-\bar{z}}{j2} \\ &\quad +\lambda\left(|z-(2+j3)|^2-2\right) \\ &= \frac{z+\bar{z}}{2}+\frac{z-\bar{z}}{j2} \\ &\quad +\lambda((z-z_0)(\bar{z}-\bar{z}_0)-2),\end{aligned} \quad (13)$$

where $z_0 \triangleq 2+j3$. Please note that the optimization problem is identical to the earlier one, given by (11). Hence, we expect the Lagrange multiplier λ to turn out to be $\lambda=\{-0.5,0.5\}$ and the final solution to be $z=1+j2$ at the end. The only difference is that the optimization is over z and $\bar{z}$, instead of (x,y).

By taking formal partial derivatives with respect to z, $\bar{z}$, and λ, we get

$$\frac{\partial}{\partial z}L(z,\bar{z},\lambda)=\frac{1}{2}+\frac{1}{j2}+\lambda(\bar{z}-\bar{z}_0)=0, \quad (14a)$$

$$\frac{\partial}{\partial \bar{z}}L(z,\bar{z},\lambda)=\frac{1}{2}-\frac{1}{j2}+\lambda(z-z_0)=0, \quad (14b)$$

$$\frac{\partial}{\partial\lambda}L(z,\bar{z},\lambda)=|z-z_0|^2-2=0. \quad (14c)$$

As before, we need to solve for z, $\bar{z}$, and λ from (14). The critical observation is that the left-hand sides of (14a) and (14b) are complex conjugates of each other and the right-hand sides of these equations are zero. Hence, if the first equation is satisfied by a particular z, then the second equation is automatically satisfied, and vice versa. Typically, readers with experience in these calculations prefer to write only one of these two equations, which is the partial derivative with respect to conjugate variable $\bar{z}$ since the other equation is redundant. The differentiation with respect to $\bar{z}$ has no real-valued calculus counterpart, thus possibly confusing the inexperienced eye.

Referring back to the problem, from (14b), we can immediately get $z=z_0-(1+j)/(2\lambda)$. As before, the Lagrange multiplier λ can be found by substituting this relation into (14c). Once this is done, we get $\lambda=\{-0.5,0.5\}$, and the extremum are at $z=z_0\mp(1+j)=\{1+j2,3+j4\}$, which coincides with our earlier findings.

The relation of conjugacy of the first two equations in (14), i.e., $\partial/\partial z\, L(z,\bar{z},\lambda)=\overline{\partial/\partial\bar{z}\, L(z,\bar{z},\lambda)}$, is not limited to the examined problem but applies to all real-valued objective functions of complex variables. This claim can immediately be verified by substituting $f(z)=u(x,y)$ into (8).

As a single sentence take-away note, we can assert the following: the conjugacy relation between partial derivatives with respect to z and $\bar{z}$ eases the calculation of the derivative, simplifies the presentation, and enables the solution of the optimization problem at the halved dimension of real-valued optimization.

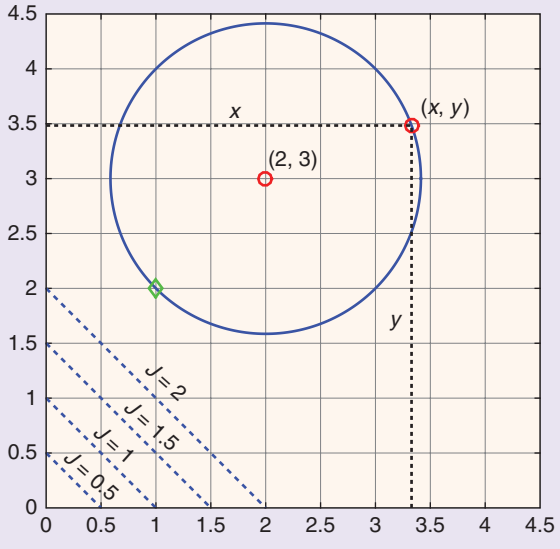

FIGURE 1. Example 1: Finding the point on the circle with the minimum coordinate sum.

Example 2: Magnitude-only least-squares filter design

The problem is to design a finite impulse-response filter whose frequency response approximates a desired magnitude response. Figure 2 shows the desired band-pass characteristic and the best least-squares approximation of the desired characteristic with a 16-coefficient filter. The filter is designed with the method described in this example. Note that the desired response in Figure 2 is not an even function of frequency; hence, the filter coefficients of this design are complex valued. The main goal of this example is to illustrate the optimization of complex-valued filter coefficients.

Let us denote the impulse response of the filter with $h[n]$, $n=\{0,\ldots,N-1\}$

and its Fourier transform, i.e., the frequency response of the system, as $H(e^{j\omega}) = \sum_{n=0}^{N-1} h[n]e^{-j\omega n}$. The design specifications for this problem are given only for the magnitude response, i.e., there are no restrictions on the phase response. Our goal is to set N complex-valued coefficients such that $|H(e^{j\omega})|^2$ approximates the desired characteristic $|H_d(e^{j\omega})|^2$ in the least-squares sense.

Consider that the magnitude square of $H(e^{j\omega})$ is also called *the energy spectral density*. This definition is only valid for the finite energy signals and is analytically expressed as $R_h(e^{j\omega}) = |H(e^{j\omega})|^2$. With this definition, the least-squares filter design problem with the magnitude response specification can be expressed as

$$\{h[0],\ldots,h[N-1]\} = \arg\min_{h[0],\ldots,h[N-1]} \int_{-\pi}^{\pi} (R_h(e^{j\omega}) - R_d(e^{j\omega}))^2 d\omega. \tag{15}$$

Here, the functions $R_h(e^{j\omega})$ and $R_d(e^{j\omega})$ are the designed and desired energy spectral density functions, respectively.

The energy spectral density function for $h[n]$ can be written as

$$R_h(e^{j\omega}) = \left|\sum_{n=0}^{N-1} h[n]e^{-j\omega n}\right|^2 = \sum_{k=-(N-1)}^{N-1} r_h[k]e^{-j\omega k}, \tag{16}$$

where $r_h[k] = \sum_{n=-\infty}^{\infty} h[n]\bar{h}[n-k]$ is the deterministic autocorrelation for the impulse-response sequence. Given that the autocorrelation sequence completely determines the energy spectral density function, the filter design problem can be decomposed into two stages: the least-squares design of the autocorrelation sequence and the filter impulse response construction from the designed autocorrelation sequence. First, we focus on the autocorrelation sequence design.

A valid autocorrelation sequence must satisfy the conjugate-symmetry property, also known as the *Hermitian symmetry property*, which is $r_h[k] = \bar{r}_h[-k]$. Hence, for a valid autocorrelation sequence, $r_h[0]$ is real valued and $r_h[k]$ is the complex conjugate of $r_h[-k]$ for $k \neq 0$. It should be noted that the Hermitian symmetry condition is just a necessary condition for the validity of the autocorrelation sequence. The necessary and sufficient condition is the nonnegativeness of its Fourier transform, i.e., the nonnegativeness of its energy spectral density or the positive definiteness of the autocorrelation sequence. From (16), the energy spectral density function can be written as

$$R_h(e^{j\omega}) = r_h[0] + \underbrace{\sum_{k=1}^{N-1} r_h[k]e^{-j\omega k}}_{S_1(e^{jw})} + \underbrace{\sum_{k=-(N-1)}^{-1} r_h[k]e^{-j\omega k}}_{S_2(e^{jw})}. \tag{17}$$

Considering the Hermitian symmetry property of $r_h[k]$, the first summation in (17) is the complex conjugate of the second summation, $S_1(e^{jw}) = \overline{S_2}(e^{jw})$. To utilize the results of vector calculus, we prefer to rewrite $S_1(e^{jw})$ as an inner product

$$S_1(e^{jw}) = [e^{-j\omega}\ e^{-j2\omega}\ \ldots\ e^{-j(N-1)\omega}] \begin{bmatrix} r_h[1] \\ r_h[2] \\ \vdots \\ r_h[N-1] \end{bmatrix} = \mathbf{v}_\omega^T \mathbf{r}, \tag{18}$$

where $\mathbf{v}_\omega = [e^{-j\omega}\ e^{-j2\omega}\ \ldots\ e^{-j(N-1)\omega}]^T$ and $\mathbf{r} = [r_h[1]\ r_h[2]\ \ldots\ r_h[N-1]]^T$. With this definition, $R_h(e^{j\omega})$, given in (17), can be formulated as

$$R_h(e^{j\omega}) = r_h[0] + S_1(e^{jw}) + S_2(e^{jw}) = r_h[0] + \mathbf{v}_\omega^T\mathbf{r} + \overline{(\mathbf{v}_\omega^T\mathbf{r})}. \tag{19}$$

Our goal is the design of an autocorrelation sequence such that the discrete-time Fourier transform of the sequence approximates the desired energy spectral density $R_d(e^{j\omega})$ in $\omega \in (-\pi, \pi]$. We can express the condition of $R_h(e^{j\omega}) \approx R_d(e^{j\omega})$ for densely populated frequency points of $\omega = \{\omega_1, \omega_2, \ldots, \omega_L\} \in (-\pi, \pi]$ as follows:

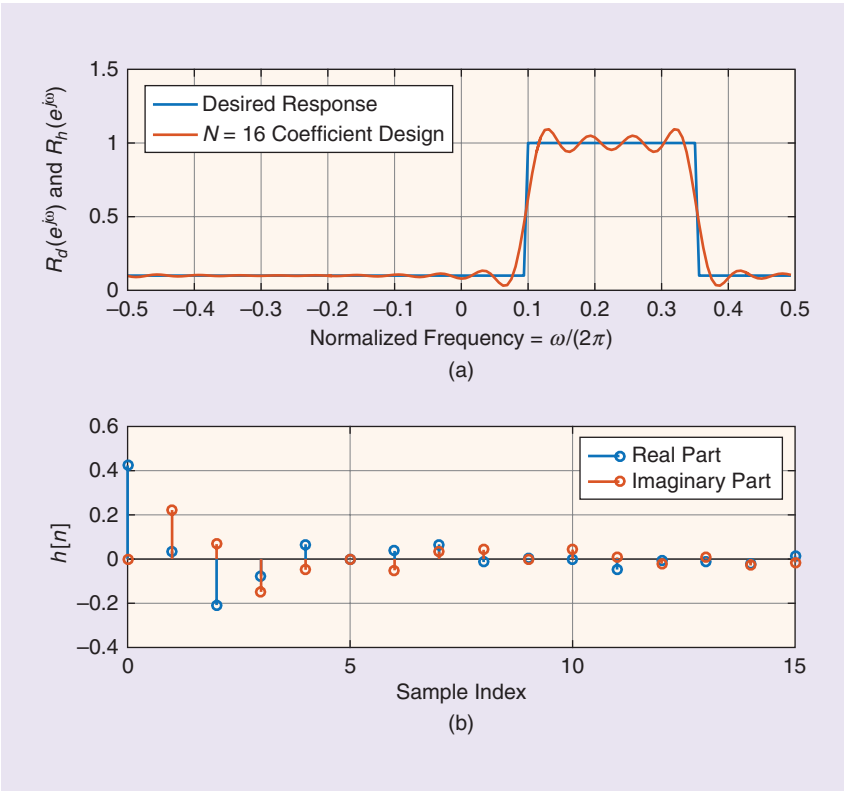

FIGURE 2. Example 2: The magnitude-only least-squares filter design. (a) The least-squares approximation to the desired magnitude response and (b) filter impulse response.

$$\underbrace{\begin{bmatrix} R_h(e^{j\omega_1}) \\ R_h(e^{j\omega_2}) \\ \vdots \\ R_h(e^{j\omega_L}) \end{bmatrix}}_{\mathbf{R_h}} = \underbrace{\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}}_{\mathbf{1}} r_h[0] + \underbrace{\begin{bmatrix} \mathbf{v}_{\omega_1}^T \\ \mathbf{v}_{\omega_2}^T \\ \vdots \\ \mathbf{v}_{\omega_L}^T \end{bmatrix}}_{\mathbf{A}} \mathbf{r} + \overline{\left(\begin{bmatrix} \mathbf{v}_{\omega_1}^T \\ \mathbf{v}_{\omega_2}^T \\ \vdots \\ \mathbf{v}_{\omega_L}^T \end{bmatrix} \mathbf{r} \right)} \approx \underbrace{\begin{bmatrix} R_d(e^{j\omega_1}) \\ R_d(e^{j\omega_2}) \\ \vdots \\ R_d(e^{j\omega_L}) \end{bmatrix}}_{\mathbf{R_d}}. \tag{20}$$

The equation system in (20) is nothing but the expression of $R_h(e^{j\omega}) \approx R_d(e^{j\omega})$ for all of the points in the frequency grid and can be compactly written in the vector-matrix form as follows.

$$\mathbf{R_h} = \mathbf{1} r_h[0] + \mathbf{A}\mathbf{r} + \bar{\mathbf{A}}\bar{\mathbf{r}} \approx \mathbf{R_d}. \tag{21}$$

We can denote the approximation error vector as $\mathbf{R_h} - \mathbf{R_d}$ and form the least-squares cost function for the approximation as follows:

$$\begin{aligned} &J(r_h[0], \ldots, r_h[N-1]) \\ &\quad = \left\| \mathbf{R_h} - \mathbf{R_d} \right\|_R^2 \\ &\quad = \left\| \mathbf{1} r_h[0] + \mathbf{A}\mathbf{r} + \bar{\mathbf{A}}\bar{\mathbf{r}} - \mathbf{R_d} \right\|_R^2. \end{aligned} \tag{22}$$

For the minimization of the cost function, the partial derivatives with respect to the real and imaginary parts of $r_h[n], n = \{0, \ldots, N-1\}$ should vanish. We prefer to evaluate the derivatives with respect to $r_h[n]$ and its conjugate $\bar{r}_h[n]$. Note that the cost function in (22) explicitly includes the unknown autocorrelation values and their conjugates. Considering that $\left\| \mathbf{x} \right\|_R^2 = \mathbf{x}^T\mathbf{x}$ for real-valued $\mathbf{x}$, the cost function can be expanded as

$$\begin{aligned} J &= \underbrace{(\mathbf{1} r_h[0] + \mathbf{A}\mathbf{r} + \bar{\mathbf{A}}\bar{\mathbf{r}} - \mathbf{R_d})^T}_{\mathbf{u}^T} \\ &\quad \underbrace{(\mathbf{1} r_h[0] + \mathbf{A}\mathbf{r} + \bar{\mathbf{A}}\bar{\mathbf{r}} - \mathbf{R_d})}_{\mathbf{u}} \\ &= \mathbf{u}^T\mathbf{u}. \end{aligned} \tag{23}$$

Since $r_h[0]$ is a real value, the partial derivative of J with respect to $r_h[0]$ easily follows from the standard calculus rules

$$\begin{aligned} \frac{\partial J}{\partial r_h[0]} &= \frac{\partial}{\partial r_h[0]} \{\mathbf{u}^T\mathbf{u}\} \\ &\overset{(a)}{=} \left(\frac{\partial \mathbf{u}^T}{\partial r_h[0]} \right) \mathbf{u} + \mathbf{u}^T \left(\frac{\partial \mathbf{u}}{\partial r_h[0]} \right) \\ &\overset{(b)}{=} \mathbf{1}^T\mathbf{u} + \mathbf{u}^T\mathbf{1} \\ &\overset{(c)}{=} 2(\mathbf{1}^T\mathbf{u}) \\ &\overset{(d)}{=} 2[\mathbf{1}^T(\mathbf{1} r_h[0] + \mathbf{A}\mathbf{r} + \bar{\mathbf{A}}\bar{\mathbf{r}} - \mathbf{R_d})]. \end{aligned} \tag{24}$$

In (24), the equality of (a) is due to the product rule for differentiation. The equality (b) stems from the fact that $r_h[0]$ is the coefficient of an all-ones vector (the vector of $\mathbf{1}$) in the expression for $\mathbf{u}$ in (23). The equalities of (c) and (d) simply follow from the definition of $\mathbf{u}$. By equating (24) to zero, we get the optimality condition for $r_h[0]$.

Next, we evaluate the gradient of the cost function with respect to $\bar{\mathbf{r}}$ to get the optimality conditions of $\bar{r}_h[k]$ for $k \neq 0$. The gradient operator is defined as $\nabla_{\bar{\mathbf{r}}} \triangleq [(\partial/\partial \bar{r}_h[1])(\partial/\partial \bar{r}_h[2]) \cdots (\partial/\partial \bar{r}_h[N-1])]^T$. The gradient of the cost function is then

$$\begin{aligned} \nabla_{\bar{\mathbf{r}}} J &= \nabla_{\bar{\mathbf{r}}} \{\mathbf{u}^T\mathbf{u}\} \\ &\overset{(a)}{=} (\nabla_{\bar{\mathbf{r}}} \mathbf{u}^T)\mathbf{u} + (\mathbf{u}^T(\nabla_{\bar{\mathbf{r}}} \mathbf{u}))^T \\ &\overset{(b)}{=} \bar{\mathbf{A}}^T\mathbf{u} + (\mathbf{u}^T\bar{\mathbf{A}})^T \\ &\overset{(c)}{=} \mathbf{A}^H\mathbf{u} + \mathbf{A}^H\mathbf{u} \\ &\overset{(d)}{=} 2\mathbf{A}^H\mathbf{u} \\ &= 2\mathbf{A}^H(\mathbf{1} r_h[0] + \mathbf{A}\mathbf{r} + \bar{\mathbf{A}}\bar{\mathbf{r}} - \mathbf{R_d}). \end{aligned} \tag{25}$$

In (25), the equality of (a) follows from the product rule for differentiation. On the right side of (a), the second term appears as transposed. This is because the gradient vector is defined as a *column vector*, and with the transpose operator on the second component, we have the addition of two column vectors. Also, $\nabla_{\bar{\mathbf{r}}} \mathbf{u}$ operation refers the calculation of the gradient for each entry of the vector $\mathbf{u}$; hence, $\nabla_{\bar{\mathbf{r}}} \mathbf{u}$ is the Jacobian matrix for the vector $\mathbf{u}$. The equality (b) stems from the fact that $\mathbf{u}^T$, defined in (23), contains a term $\bar{\mathbf{r}}^T\bar{\mathbf{A}}^T$ and $\nabla_{\bar{\mathbf{r}}} \{\bar{\mathbf{r}}^T\bar{\mathbf{A}}^T\} = \bar{\mathbf{A}}^T$. The equality (c) is due to the Hermitian operation definition, i.e., $\mathbf{A}^H = \bar{\mathbf{A}}^T$. Finally, by equating (25) to an all-zero vector, we get the optimality conditions for the variation of $\bar{r}_h[k]$ for $k \neq 0$.

The optimality conditions for the variation of $r_h[k]$ for $k \neq 0$ can also be derived by calculating the gradient with respect to $r_h[k]$, which is a procedure very similar to the one given in (25). Since the cost function is a real-valued function of complex variables, we can avoid this calculation, as in Example 1. As discussed before, this feature of the cost function leads to the fact that the partial derivatives with respect to the variable and its conjugate are complex conjugates of each other, $\nabla_{\mathbf{r}} J = \overline{(\nabla_{\bar{\mathbf{r}}} J)}$. Using this fact, the optimality condition of $r_h[k]$ for $k \neq 0$ can be immediately written from the conditions for the conjugate variable

$$\begin{aligned} \nabla_{\mathbf{r}} J = \overline{(\nabla_{\bar{\mathbf{r}}} J)} &\overset{(25d)}{=} \overline{(2\mathbf{A}^H\mathbf{u})} = 2\mathbf{A}^T\mathbf{u} \\ &= 2\mathbf{A}^T(\mathbf{1} r_h[0] + \mathbf{A}\mathbf{r} + \bar{\mathbf{A}}\bar{\mathbf{r}} - \mathbf{R_d}). \end{aligned} \tag{26}$$

By combining the optimality conditions for $r_h[0], r_h[k]$, and $\bar{r}_h[k]$ and equating the partial derivatives given by (24)–(26) to zero, we get the following linear equation system:

$$\begin{bmatrix} \mathbf{1}^T\mathbf{1} & \mathbf{1}^T\mathbf{A} & \mathbf{1}^T\bar{\mathbf{A}} \\ \mathbf{A}^H\mathbf{1} & \mathbf{A}^H\mathbf{A} & \mathbf{A}^H\bar{\mathbf{A}} \\ \mathbf{A}^T\mathbf{1} & \mathbf{A}^T\mathbf{A} & \mathbf{A}^T\bar{\mathbf{A}} \end{bmatrix} \begin{bmatrix} r_h[0] \\ \mathbf{r} \\ \bar{\mathbf{r}} \end{bmatrix} = \begin{bmatrix} \mathbf{1}^T \\ \mathbf{A}^H \\ \mathbf{A}^T \end{bmatrix} \mathbf{R_d}. \tag{27}$$

The solution of the equation system in (27) gives the optimal design for the autocorrelation sequence in the least-squares sense. The best least-squares approximation to the desired energy spectral density is illustrated for a 16-coefficient filter in Figure 2. As can be seen from the top panel of Figure 2, the approximation is nonnegative valued, i.e., it satisfies the necessary and sufficient condition on a valid autocorrelation sequence. The design suffers from oscillatory behavior which, in general, is typical for the least-squares problems.

In the second stage, the filter impulse response is calculated from its autocorrelation sequence. Since $r_h[k] = \sum_{n=-\infty}^{\infty} h[n]\bar{h}[n-k]$, the impulse-response $h[n]$ can be retrieved from its autocorrelation via spectral factorization, [4]. The spectral factorization result for a minimum-phase $h[n]$ is shown in Figure 2. We do not provide any more details on the implementation of the spectral factorization operation, which uses somewhat more advanced topics in the theory of complex-valued functions but interested readers can examine and experiment with ready-to-use MATLAB code given in [5] for more information.

Examples 1 and 2 are given to illustrate the application of Wirtinger calculus in the optimization problems. Wirtinger calculus can also be utilized in the Taylor series expansion, i.e., in the approximation problems involving functions of complex variables.

Taylor series expansion of real-valued functions of complex variables

In many applications, nonlinear relations involving complex variables are approximated around a suitable operating point with a few terms of Taylor series expansion. Such approximations are not only inevitable in the numerical optimization routines but also important in theoretical developments, such as the Cramer–Rao bound calculations, the maximum-likelihood estimation of parameters, and so on.

We illustrate the expansion process with the $f(z) = \log(1 - |z|^2)$ function. First, we express $f(z)$ as a function of z and $\bar{z}$, $f(z, \bar{z}) = \log(1 - z\bar{z})$. Treating z and $\bar{z}$ as independent variables and expanding the function of two variables into the Taylor series, we get

$$f(z) = f(z_0, \bar{z}_0) + \underbrace{\begin{bmatrix} f_z & f_{\bar{z}} \end{bmatrix}}_{\text{Gradient at } z=z_0} \begin{bmatrix} z - z_0 \\ \bar{z} - \bar{z}_0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} z - z_0 & \bar{z} - \bar{z}_0 \end{bmatrix} \underbrace{\begin{bmatrix} f_{zz} & f_{z\bar{z}} \\ f_{\bar{z}z} & f_{\bar{z}\bar{z}} \end{bmatrix}}_{\text{Hessian at } z=z_0} \begin{bmatrix} z - z_0 \\ \bar{z} - \bar{z}_0 \end{bmatrix} + \text{h.o.t.}, \quad (28)$$

where h.o.t refers to the *higher-order terms of the Taylor series expansion*. The functions $f(\cdot)$ with the subscript of z or $\bar{z}$ appearing at the gradient vector and Hessian matrix in (28) denote the evaluation of gradient and Hessian at the operating point of $z = z_0$. For example, $f_{z\bar{z}}$ refers to the numerical value of $f_{z\bar{z}} = \partial^2 / \partial z \partial \bar{z}\, f(z_0, \bar{z}_0)$.

We now see the structure imposed by real-valued functions of complex variables on the gradient vector and the Hessian matrix in (28). The first and second elements of the gradient vector are conjugates of each other, since $f_z = \overline{f_{\bar{z}}}$. (Note that this is only valid for the real-valued $f(z)$ functions.) An approximation of keeping the first two terms of (28), i.e., constant and linear terms, result in a real-valued function, as expected.

To determine the structure of the Hessian matrix in (28), we rewrite the fundamental relations in (8) with the operator notation as

$$\begin{aligned} \frac{\partial}{\partial z} f(z, \bar{z}) &\overset{\text{eq.(8)}}{=} \frac{1}{2}\left(\frac{\partial f(z)}{\partial x} - j \frac{\partial f(z)}{\partial y} \right) \\ &= D_z\{f(z)\} \\ \frac{\partial}{\partial \bar{z}} f(z, \bar{z}) &\overset{\text{eq.(8)}}{=} \frac{1}{2}\left(\frac{\partial f(z)}{\partial x} + j \frac{\partial f(z)}{\partial y} \right) \\ &= D_{\bar{z}}\{f(z)\}. \end{aligned} \quad (29)$$

where operators of D_z and $D_{\bar{z}}$ are formally defined as $D_z = (1/2)((\partial/\partial x) - j(\partial/\partial y))$ and $D_{\bar{z}} = (1/2)((\partial/\partial x) + j(\partial/\partial y))$. With the operator definitions, partial derivatives for all of the orders can be written as

$$\begin{aligned} \frac{\partial^{k+l}}{\partial z^k \partial \bar{z}^l} f(z, \bar{z}) &= D_z^k\{D_{\bar{z}}^l\{f(z, \bar{z})\}\} \\ &= D_z^k D_{\bar{z}}^l\{f(z, \bar{z})\}. \end{aligned} \quad (30)$$

In passing to the rightmost equality of (30), the outer curly brackets in the middle term are removed due to the commutativity of the operators D_z and $D_{\bar{z}}$. Also considering the fact that $D_z = \overline{D_{\bar{z}}}$, we reach the following conclusions for the entries of Hessian matrix: $f_{zz} = \overline{f_{\bar{z}\bar{z}}}$ and $f_{z\bar{z}} = f_{\bar{z}z}$. Hence, the Hessian matrix in (28) can be characterized as a *centro-Hermitian symmetric matrix*. This concludes our discussion on optimization and approximation by the Taylor series for real-valued functions of complex variables.

As an example, the Taylor series expansion of $f(z) = \log(1 - |z|^2)$, or equivalently $f(z, \bar{z}) = \log(1 - z\bar{z})$ at $z_0 = j/\sqrt{2}$, can be written as

$$\begin{aligned} f(z) = {} & f\left(\frac{j}{\sqrt{2}}, \frac{-j}{\sqrt{2}} \right) \\ & + \sqrt{2}\begin{bmatrix} j & -j \end{bmatrix} \begin{bmatrix} z - j/2 \\ \bar{z} + j/2 \end{bmatrix} \\ & + \frac{1}{2}\begin{bmatrix} z - j/2 \\ \bar{z} + j/2 \end{bmatrix}^T \begin{bmatrix} 2 & -4 \\ -4 & 2 \end{bmatrix} \begin{bmatrix} z - j/2 \\ \bar{z} + j/2 \end{bmatrix} \\ & + \text{h.o.t.} \end{aligned} \quad (31)$$

The explicit expressions for the partial derivatives are

$$\begin{aligned} f_z(z, \bar{z}) &= \frac{\partial}{\partial z} f(z, \bar{z}) = -\frac{\bar{z}}{1 - z\bar{z}}, \\ f_{zz}(z, \bar{z}) &= \frac{\partial^2}{\partial z^2} f(z, \bar{z}) = -\frac{\bar{z}^2}{(1 - z\bar{z})^2}, \\ f_{z\bar{z}}(z, \bar{z}) &= \frac{\partial^2}{\partial z \partial \bar{z}} f(z, \bar{z}) = -\frac{1}{(1 - z\bar{z})^2}. \end{aligned}$$

The first- and second-order derivatives obey the symmetry relations $f_z = \bar{f}_{\bar{z}}$ and $f_{zz} = \bar{f}_{\bar{z}\bar{z}}$, $f_{z\bar{z}} = f_{\bar{z}z}$, as anticipated.

Case 2: Complex-valued functions of complex variables

Complex-valued functions from $z \in \mathbb{C}$ to $f(z) \in \mathbb{C}$, such as $f(z) = z^2$, cannot be utilized for optimization purposes due to the basic fact that complex numbers cannot be ordered. Stated differently, $1 + j$ cannot be compared with $2 + j3$; only the magnitude of complex numbers can be compared, as in $|2 + j3| > |1 + j|$. In signal processing, complex-valued functions establish mappings between complex-valued entities, such as the mapping between filter coefficients and pole/zero locations. In this section, we focus on the Taylor series-based approximations for such functions.

Complex-valued nonanalytic functions

Lattice filters are known to have certain implementation advantages over classical structures, such as robustness to the coefficient quantization errors, order-recursive structure, and ease of stability check, and so on [4]. Assuming an all-pole filter with the transfer function of $H(z) = 1/(1 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3})$, which is implemented with the reflection coefficients $\Gamma_1, \Gamma_2, \Gamma_3$, the mapping between filter coefficients and reflection coefficients, known as *step-up recursion*, can be written as follows [4, p. 234]:

$$\begin{aligned} \mathbf{a}(\Gamma_1, \Gamma_2, \Gamma_3) &= \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \\ &= \begin{bmatrix} \Gamma_1 + \bar{\Gamma}_1 \Gamma_2 + \Gamma_3 \bar{\Gamma}_2 \\ \Gamma_2 + \Gamma_3 \bar{\Gamma}_1 + \Gamma_3 \Gamma_1 \bar{\Gamma}_2 \\ \Gamma_3 \end{bmatrix}. \end{aligned} \quad (32)$$

Here, $\mathbf{a}(\cdot)$ is a complex-valued vector function of complex variables. The second entry of this vector-valued function, $a_2(\Gamma_1, \Gamma_2, \Gamma_3) \triangleq \Gamma_2 + \Gamma_3 \bar{\Gamma}_1 + \Gamma_3 \Gamma_1 \bar{\Gamma}_2$, is readily expressed in terms of the reflection coefficients and their conjugates. Considering our earlier discussions on the Cauchy–Riemann conditions, we can immediately note that the function $a_2(\cdot)$ is not an analytic function in the complex analysis sense because of its dependence on the conjugate variables. The lack of analyticity property is not critically important in many signal processing applications, except the ones involving path integration, i.e., the residue calculus. Here, we focus on

the Taylor series expansion of nonanalytic functions such as $a_2(\cdot)$.

The partial derivative of $a_2(\cdot)$ with respect to $\bar{\Gamma}_2$ is $(\partial a_2/\partial\bar{\Gamma}_2) = \Gamma_3\Gamma_1$, a fact that can be immediately verified upon inspection of (32). Unlike the real-valued functions of complex variables previously examined, the partial derivative with respect to a variable and its conjugate do not carry information about each other for the complex-valued functions. Continuing with this current example, the derivative of $a_2(\cdot)$ with respect to Γ_2 can be noted as $(\partial a_2/\partial\Gamma_2) = 1$, which is clearly not exhibiting any connections with $(\partial a_2/\partial\bar{\Gamma}_2) = \Gamma_3\Gamma_1$. In summary, the Taylor series expansion of complex-valued functions can be written in the form given in (28), with the difference that there is no structure in the gradient and Hessian terms for such expansions.

Complex-valued analytic functions

If the complex-valued function $f(\cdot)$ is analytic, then, as previously noted, the function solely depends on z but not on its conjugate $\bar{z}$ in a domain. For such functions, the Taylor series, as in (28), can be greatly simplified. Because the derivatives involving conjugate variable $\bar{z}$ simply vanish, $f(z)$ can be expanded into the Taylor series as a function of the single, independent complex variable

$$f(z) = f(z_0) + f_z(z_0)(z - z_0) + f_{zz}(z_0)\frac{(z - z_0)^2}{2} + \text{h.o.t.} \quad (33)$$

Both (28) and (33) are Taylor series expansions for a function of a single, complex variable. Yet, they are very different from each other in appearance because of the difference in their analyticity properties. This situation may potentially confuse a novice researcher.

Nearly all of the elementary functions of real-valued calculus such as $x^k, \sin(x), \exp(x)$, and so on can be extended to the complex plane by replacing the real-valued x variable with a complex-valued z. This simple substitution has important consequences. Extended (i.e., analytically continued) functions in the complex plane attain several important features, such as infinite-order differentiation, closed contour integration via residue calculus, and many more. Looking from a broader perspective, we can consider the elementary functions of real variables as the functions with pruned features because of the domain restriction on the real line.

For more information on analytic functions and other topics, the book by Churchill and Brown [3] is the standard applied mathematics textbook on complex variables. With its approachable style, without sacrificing rigor, this book is highly recommended for all readers. Along with the Churchill's book, interested readers can also check the online lecture notes of Terence Tao on complex analysis [6]. Readers interested in the history of complex analysis will certainly enjoy the informative tour of all the major topics of complex analysis guided by P. Nahin [7]. Readers with more knowledge on the topics should definitely make time to watch V. Balakrishnan's lectures on complex analysis [8]. Finally, complex differentiation and other related issues discussed in these notes are the basic tools of the research area known as *widely linear estimation theory*. For research applications, readers can consult [2], [9], and [10].

What we have learned

We have studied the definition of complex differentiation and its implications in optimization and approximation problems. We have seen that a real-valued function of complex variables, which is basically a cost function, can be optimized with a reduced effort at halved dimension of real-valued optimization by complex differentiation with respect to the conjugate variable $\bar{z}$. We have studied approximation by the Taylor series and noted certain differences for analytic and nonanalytic functions. Most of the elementary functions of standard calculus remain as analytic functions with the replacement of x with the complex variable z. For analytic functions, the Taylor series expansion for real variables and complex variables are identical in form; however, this is not the case in general.

Acknowledgments

I would like to thank Prof. Rodrigo Capobianco Guido of São Paulo State University, Brazil, and Prof. Umut Orguner of Middle East Technical University, Ankara, Turkey, for their kind help, many suggestions, and insightful comments.

Author

Çağatay Candan (ccandan@metu.edu.tr) received his B.S., M.S., and Ph.D. degrees, all in electrical engineering, from Middle East Technical University, Ankara, Turkey, Bilkent University, Ankara, and the Georgia Institute of Technology, Atlanta, in 1996, 1998, and 2004, respectively. Currently, he is a professor in the Department of Electrical and Electronics Engineering at Middle East Technical University, where he assists students in establishing a solid engineering foundation upon which upcoming applications and future careers can be built. His research interests include statistical signal processing and its applications in array signal processing and radar/sonar signal processing.

References

[1] W. Wirtinger, "Zur formalen Theorie der Funktionen von mehr komplexen Veränderlichen," *Mathematische Annalen*, vol. 97, no. 1, pp. 357–375, 1927.

[2] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data*. Cambridge, U.K.: Cambridge Univ. Press, 2010.

[3] J. W. Brown and R. V. Churchill, *Complex Variables and Applications*, 8th ed. New York: McGraw-Hill, 2008.

[4] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*. New York: Wiley, 1996.

[5] C. Candan. (2018). Magnitude-only filter design (MATLAB codes). Middle East Technical University, Ankara, Turkey. [Online]. Available: http://users.metu.edu.tr/ccandan/matlab/SPMfiltdesign

[6] T. Tao. (2016). Academic blog: 246A complex analysis. [Online]. Available: https://terrytao.wordpress.com/category/teaching/246a-complex-analysis

[7] P. J. Nahin, *An Imaginary Tale: The Story of* $\sqrt{-1}$. Princeton, NJ: Princeton Univ. Press, 2007.

[8] V. Balakrishnan. Selected topics in mathematical physics. Ministry of Human Resource Development, India. [Online]. Available: http://nptel.ac.in/courses/115106086/

[9] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 2030–2033, 1995.

[10] T. Adali and P. J. Schreier, "Optimization and estimation of complex-valued signals: Theory and applications in filtering and blind source separation," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 112–128, 2014.

SP