Herglotz functions and applications in electromagnetics

Nedic, Mitja

Institution of engineering and technology
2021-01-01

Nedic , M , Ehrenborg , C , Ivanenko , Y , Ludvig-Osipov , A , Nordebo , S , Luger , A , Jonsson , L , Sjöberg , D & Gustafsson , M 2021 , Herglotz functions and applications in electromagnetics . in K Kobayashi & P D Smith (eds) , Advances in Mathematical Methods for Electromagnetics . Institution of engineering and technology , pp. 491-514 . https://doi.org/10.1049/SBEW528E_c

http://hdl.handle.net/10138/355593
https://doi.org/10.1049/SBEW528E_ch20

unspecified
acceptedVersion

Chapter 20

Herglotz functions and applications in electromagnetics

Mitja Nedic[1], Casimir Ehrenborg[2], Yevhen Ivanenko[3], Andrei Ludvig-Osipov[4], Sven Nordebo[5], Annemarie Luger[6], Lars Jonsson[7], Daniel Sjöberg[8], and Mats Gustafsson[9]

Herglotz functions inevitably appear in pure mathematics, mathematical physics, and engineering with a wide range of applications. In particular, they are the pertinent functions to model passive systems, and thus appear in modeling of electromagnetic phenomena in circuits, antennas, materials, and scattering. In this chapter, we review the basic theory of Herglotz functions and its applications to determine sum rules and physical bounds for passive systems.

20.1 Introduction

Holomorphic mappings between certain half-planes appear in areas such as spectral theory [1,2], moment problems [3,4], passive systems [5,6], circuit synthesis [7,8], dispersion relations [9,10], and homogenization [11,12]. These functions, here referred to as Herglotz functions, are also known as Nevanlinna, Herglotz-Nevanlinna, Pick, R- [13], and positive real (PR) [7] functions. In this chapter, we review mathematical properties of Herglotz functions and discuss their applications in the modeling of passive systems, derivation of sum rules, and physical bounds.

Herglotz functions can be represented by an integral representation solely depending on scalar parameters and a positive measure. This representation is very powerful and is the starting point for many results on Herglotz functions. In particular, we use the integral representation to derive identities that relate weighted integrals of a Herglotz function with its asymptotic expansion [3]. These identities are often referred to as sum rules and have many applications in electromagnetics [14].

[1]Department of Mathematics and Statistics, University of Helsinki, Helsinki, Finland
[2]Department of Electrical Engineering, Katholieke Universiteit Leuven, Leuven, Belgium
[3]Department of Mathematics and Natural Sciences, Blekinge Institute of Technology, Karlskrona, Sweden
[4]Department of Electrical Engineering, Chalmers University of Technology, Göteborg, Sweden
[5]Department of Physics and Electrical Engineering, Linnaeus University, Växjö, Sweden
[6]Department of Mathematics, Stockholm University, Stockholm, Sweden
[7]School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden
[8]Department of Electrical and Information Technology, Lund University, Lund, Sweden

Passivity is instrumental in the modeling of electromagnetic phenomena. It is used to describe objects and systems that do not produce energy. Here, a system perspective is employed where passivity characterizes systems where the energy leaving the system does not exceed the energy that has entered the system for all signals and times. Linear, time-translational invariant, and passive systems have transfer functions that either are or can be transformed to Herglotz and PR functions [6]. This implies that Herglotz functions constitute a generic way to model all passive systems regardless of the complexity of the system.

Sum rules are useful in many branches of physics and engineering as they relate dynamic parameter values with their low- and high-frequency expansions [9]. Properties of the dynamic response can hence be inferred by the, in many cases much simpler, static response. These types of identities and physical bounds are of great interest in many areas of electromagnetic theory. They provide physical insight of the relation between design parameters and are useful in optimization as they provide upper bounds on the design. The approach is based on the identification of a passive system where a parameter of interest is the imaginary part of a Herglotz function. This has, e.g., been used for lossless matching networks [15], radar absorbers [16], extinction cross section [17], antennas [18], high-impedance surfaces [19,20], metamaterials [21], and array antennas [22,23].

In this chapter, we show that the integral identities for Herglotz functions constitute a unified approach to derive sum rules and illustrate the approach for the input impedance of lumped circuit networks, temporal dispersion of metamaterials, and radar absorbers.

Convex optimization can be used together or as an alternative to sum rules to determine physical bounds [24,25]. This greatly extends the class of solvable problems for the price of numerical solutions. This approach is based on the linearity and positivity of the measure in the integral representation implying that the set of Herglotz functions is a convex cone. Convex optimization can also be used to characterize rational PR functions and passive systems via the PR Lemma [5,26] and in conjunction with Nevanlinna–Pick interpolation [27].

This chapter starts with a review of the definitions, basic properties, and integral identities of Herglotz functions in Section 20.2. Passive systems in electromagnetics and their relation to Herglotz functions are discussed in Section 20.3. Sum rules and physical bounds are investigated in Section 20.4. The application of Herglotz functions in convex optimization for obtaining physical bounds is reviewed in Section 20.5.

20.2 Basics about Herglotz functions

In this section, we introduce some standard notations and review some classical results about Herglotz functions. Throughout this text, we denote the open upper and right half-planes, respectively, by

$$\mathbb{C}^+ := \{z \in \mathbb{C} \mid \mathrm{Im}[z] > 0\} \quad \text{and} \quad \mathbb{C}_+ := \{z \in \mathbb{C} \mid \mathrm{Re}[z] > 0\} \tag{20.1}$$

The open lower and left half-planes are defined analogously and are denoted by $\mathbb{C}^-$ and $\mathbb{C}_-$, respectively (see Figure 20.1). In connection with these particular subsets

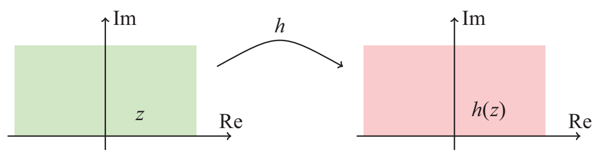


Figure 20.1 A Herglotz function maps the upper half-plane $\mathbb{C}^+$ (green) to the closed upper half-plane $\mathbb{C}^+ \cup \mathbb{R}$ (red)

of the complex plane $\mathbb{C}$, we are interested in the following classes of holomorphic (cf. [28]) functions.

Definition 20.1. A function $h\colon \mathbb{C}^+ \to \mathbb{C}$ is called a *Herglotz function* if it is holomorphic with $\mathrm{Im}[h(z)] \geq 0$ for all $z \in \mathbb{C}^+$. □

Example 20.1. Some basic examples of Herglotz functions are the following:

$$h_1(z) = 1, \quad h_2(z) = z, \quad h_3(z) = \mathrm{i}, \quad \text{and} \quad h_4(z) = -\frac{1}{z} \tag{20.2}$$

A few less trivial Herglotz functions (cf. [29]) are

$$h_5(z) = \tan(z), \quad h_6(z) = -\cot(z), \quad \text{and } h_7(z) = \frac{\Gamma'(z)}{\Gamma(z)} \tag{20.3}$$

where Γ denotes Euler's gamma function (cf. [30]) and Γ' is its derivative. ■

New Herglotz functions can be constructed using the following properties:

a. The set of all Herglotz functions is a convex cone; that is, any positive linear combination of Herglotz functions is again a Herglotz function.
b. For any two Herglotz functions h_{I} and h_{II}, such that the function h_{I} does not attain a real value, we can form the composition $z \mapsto h_{\mathrm{II}}(h_{\mathrm{I}}(z))$, which is again a Herglotz function.

For example, using only the known Herglotz functions from Example 20.1 and the two properties (a) and (b) stated previously, we can conclude that the function

$$z \mapsto 1 + z - \frac{1}{2z - \frac{3}{z} + 4\mathrm{i}} \tag{20.4}$$

is also a Herglotz function.

A common special case is to consider the composition of a Herglotz function with the function $z \mapsto \mathrm{Log}(z)$, as in the following example.

Example 20.2. First, let Log be the function defined as

$$\mathrm{Log}\colon z \mapsto \ln|z| + \mathrm{i}\,\mathrm{Arg}(z) \tag{20.5}$$

where the argument of complex number is taken from the interval $[-\pi, \pi)$. Observe that with this definition, the complex logarithm becomes a Herglotz function.

Let now h be a Herglotz function such that $h(z) \notin (-\infty, 0]$ for any $z \in \mathbb{C}^+$. The function $z \mapsto \mathrm{Log}(h(z))$ is therefore well-defined and, in fact, is also a Herglotz function, since $\mathrm{Im}[\mathrm{Log}(h(z))] = \mathrm{Arg}(h(z)) \geq 0$. ■

An important subclass of Herglotz functions is given in the following definition.

Definition 20.2. A Herglotz function h, satisfying the additional condition that

$$h(-z^*) = -h(z)^* \tag{20.6}$$

is called a *symmetric Herglotz function*. □

Herglotz functions appear, for instance, in quantum mechanics, whereas in system theory, as well as in other places, also the following class of functions is widely used.

Definition 20.3. A holomorphic function $p\colon \mathbb{C}_+ \to \mathbb{C}$ is called positive real (PR) if it has $\mathrm{Re}[p(z)] \geq 0$ for $z \in \mathbb{C}_+$ and takes real values on the positive real line (Figure 20.2). □

Any symmetric Herglotz function h gives rise to a PR function p by setting $p(z) := -\mathrm{i}\, h(\mathrm{i}\, z)$ and, conversely, any PR function p gives rise to a symmetric Herglotz function h by setting $h(z) := \mathrm{i}\, p(-\mathrm{i}\, z)$.

One of the most powerful tools in the theory of Herglotz functions is the classical integral representation theorem. The theorem was first considered over 100 years ago and has paved the way for many other results about Herglotz functions. Using modern notation, we present the theorem in the following form (see also [13]).

Theorem 20.1. A function $h\colon \mathbb{C}^+ \to \mathbb{C}$ is a Herglotz function if and only if h can be written as

$$h(z) = a + b\,z + \int_{\mathbb{R}} \left(\frac{1}{\tau - z} - \frac{\tau}{1+\tau^2} \right) \mathrm{d}\mu(\tau) \tag{20.7}$$

where μ is a positive Borel measure (cf. [32]) satisfying the growth condition

$$\int_{\mathbb{R}} \frac{1}{1+\tau^2}\, \mathrm{d}\mu(\tau) < \infty \tag{20.8}$$

and the constants $a \in \mathbb{R}$ and $b \geq 0$. □

The following properties can be derived from the above theorem [13,31]:

i. If a Herglotz function attains a real value in the open upper half-plane, then it is a real-constant function.

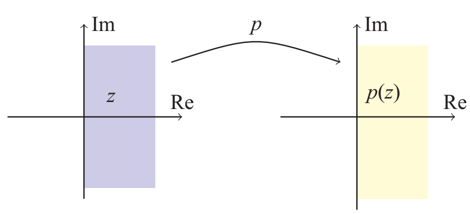


Figure 20.2 A PR function maps the right half-plane $\mathbb{C}_+$ (blue) to the closed right half-plane $\mathbb{C}_+ \cup \mathrm{i}\mathbb{R}$ (yellow) and, in particular, $\mathbb{R}^+ := (0,\infty)$ into $\mathbb{R}_0^+ := [0,\infty)$

ii. The number a from Theorem 20.1 is equal to $a = \mathrm{Re}[h(\mathrm{i})]$ while the number b is given by

$$b = \lim_{z \hat{\to} \infty} \frac{h(z)}{z} \tag{20.9}$$

Here, the symbol $\hat{\to}$ denotes a non-tangential limit (see Remark 20.1).

iii. The measure μ from Theorem 20.1 is given as the distributional boundary value of $\mathrm{Im}[h]$. More precisely, for any C^1- (i.e., continuously differentiable) function $\varphi\colon \mathbb{R} \to \mathbb{R}$ which satisfies the inequality $|\varphi(x)| \leq C(1+x^2)^{-1}$ for some $C \geq 0$ and all $x \in \mathbb{R}$, we have that

$$\lim_{y \to 0^+} \frac{1}{\pi} \int_{\mathbb{R}} \varphi(x) \mathrm{Im}[h(x + \mathrm{i}\, y)]\, \mathrm{d}x = \int_{\mathbb{R}} \varphi(\tau)\, \mathrm{d}\mu(\tau) \tag{20.10}$$

iv. The measure μ has a point mass at the point $\tau_0 \in \mathbb{R}$ if and only if the limit

$$\mu(\{\tau_0\}) = \lim_{z \hat{\to} \tau_0} (\tau_0 - z) h(z) \tag{20.11}$$

is positive, [13].

v. If the measure μ is purely absolutely continuous with respect to the Lebesgue measure $\lambda_{\mathbb{R}}$ and the boundary value $\mu'(x) := \lim_{y \to 0^+} \frac{1}{\pi} \mathrm{Im}[h(x + \mathrm{i}y)]$ exists as a square integrable function, i.e., $\mu' \in \mathrm{L}^2(\mathbb{R})$, then

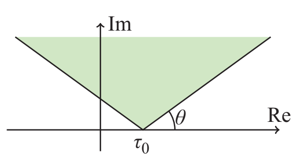
$$h(z) = \tilde{a} + b\, z + \int_{\mathbb{R}} \frac{\mu'(\tau)}{\tau - z}\, \mathrm{d}\tau, \quad \text{where } \tilde{a} := a - \int_{\mathbb{R}} \frac{\tau\, \mu'(\tau)}{1 + \tau^2}\, \mathrm{d}\tau \tag{20.12}$$

vi. Symmetric Herglotz functions (20.6) admit the integral representation

$$h(z) = b\, z + \lim_{R \to \infty} \int_{-R}^{R} \frac{\mathrm{d}\mu(\tau)}{\tau - z} = b\, z - \frac{c}{z} + \int_{(0,\infty)} \frac{2z}{\tau^2 - z^2}\, \mathrm{d}\mu(\tau) \tag{20.13}$$

where $c = \mu(\{0\})$.

Remark 20.1. The notation $z \hat{\to} \infty$ denotes the non-tangential limit $|z| \to \infty$, i.e., within some *Stoltz domain* $\{z \in \mathbb{C}^+ \mid \theta \leq \mathrm{Arg}(z) \leq \pi - \theta\}$ with parameter $\theta \in (0, \frac{\pi}{2}]$. Similarly, $z \hat{\to} \tau_0$ denotes the non-tangential limit $z \to \tau_0$ within $\{z \in \mathbb{C}^+ \mid \theta \leq \mathrm{Arg}(z - \tau_0) \leq \pi - \theta\}$ with $\theta \in (0, \frac{\pi}{2}]$. A Stoltz domain in the latter case is visualized in the figure below.



The integral representation offers also a way to synthesize Herglotz functions.

Example 20.3. The point measure δ_{x_0} at a point $x_0 \in \mathbb{R}$, along with suitably chosen constants a and b, produces the rational Herglotz function $h(z) = 1/(x_0 - z)$ with a simple pole at the point x_0; see also (20.11). Here, we note that $\mathrm{Im}[h(x)] = 0$ for

$x \in \mathbb{R} \setminus \{x_0\}$ while $h(x_0)$ is undefined. Similarly, the sum of the point masses $\delta_{(n-\frac{1}{2})\pi}$ at the points $(n - \frac{1}{2})\pi$, $n \in \mathbb{Z}$ produces the tangent function, which is thus a Herglotz function, i.e.,

$$\tan(z) = \sum_{n=-\infty}^{\infty} \left(\frac{1}{(n-\frac{1}{2})\pi - z} - \frac{(n-\frac{1}{2})\pi}{1 + (n-\frac{1}{2})^2\pi^2} \right) = \sum_{n=1}^{\infty} \frac{8z}{(2n-1)^2\pi^2 - 4z^2} \tag{20.14}$$

where the first expression includes the convergence term in (20.7) and the symmetric version (20.13) is used in the second expression. ■

The class of Herglotz functions has a non-empty intersection with Hardy spaces (cf. [33]). The Hardy space H^p, where $1 \leq p \leq \infty$, is the space of holomorphic functions in $\mathbb{C}^+$ for which the norm

$$\|g\|_{\mathrm{H}^p} := \begin{cases} \sup\limits_{y>0} \left(\int_{\mathbb{R}} |g(x + \mathrm{i}y)|^p \,\mathrm{d}x \right)^{\frac{1}{p}} & 1 \leq p < \infty \\ \sup\limits_{z \in \mathbb{C}^+} |g(z)| & p = \infty \end{cases} \tag{20.15}$$

is finite. In particular, for a function $g \in \mathrm{H}^p$ for almost all $x \in \mathbb{R}$ the boundary value $\hat{g}(x) := \lim_{y \to 0^+} g(x + \mathrm{i}y)$ exist and $\hat{g} \in \mathrm{L}^p(\mathbb{R})$.

Hardy spaces are also related to the Hilbert transform [9]. The Hilbert transform of a function u is defined as

$$\mathscr{H}\{u\}(x) := \frac{1}{\pi} \,\text{p.v.} \int_{\mathbb{R}} \frac{u(\tau)}{x - \tau} \,\mathrm{d}\tau \tag{20.16}$$

where p.v. means the Cauchy principal value of the integral. The operator $\mathscr{H}$ is bounded in L^p for $1 < p < \infty$; see [9] for a discussion of the L^1 and L^∞ cases.

The real and imaginary parts of the boundary value of a function $g \in \mathrm{H}^p$ are related by the Hilbert transform, i.e., for $u(x) := \mathrm{Re}[\hat{g}(x)]$ and $v(x) := \mathrm{Im}[\hat{g}(x)]$ it holds that

$$u = -\mathscr{H}\{v\} \quad \text{and} \quad v = \mathscr{H}\{u\} \tag{20.17}$$

Functions $g \in \mathrm{H}^p$ with $\mathrm{Im}[g(z)] \geq 0$ are Herglotz functions and the corresponding Hilbert transform relations are often referred to as the Kramers–Kronig relations or dispersion relations [10]. On the contrary, some Herglotz functions do not belong to any space H^p, some examples being the functions h_2, h_4, h_5, h_6, and h_7 from Example 20.1.

Another feature of Herglotz functions is that they, under certain conditions, provide integral identities closely connected to the moments of the measures with interesting applications to electromagnetic theory. Let us begin with some necessary definitions.

Definition 20.4. Let h be a Herglotz function. If for $K \geq -1$, there exist real numbers $b_1, b_0, b_{-1}, \ldots, b_{-K}$ such that h can be written as

$$h(z) = b_1 z + b_0 + \frac{b_{-1}}{z} + \cdots + \frac{b_{-K}}{z^K} + o\left(\frac{1}{z^K}\right) \qquad \text{as } z \hat{\to} \infty \tag{20.18}$$

then we say that h admits at $z = \infty$ an *asymptotic expansion of order K*. □

Remark 20.2. This means that

$$\lim_{z \hat{\to} \infty} z^K \left(h(z) - b_1 z - b_0 - \frac{b_{-1}}{z} - \cdots - \frac{b_{-K}}{z^K} \right) = 0 \tag{20.19}$$

Moreover, the coefficients b_{-j} are given by

$$b_{-j} = \lim_{z \hat{\to} \infty} z^j \left(h(z) - b_1 z - b_0 - \frac{b_{-1}}{z} - \cdots - \frac{b_{-(j-1)}}{z^{j-1}} \right) \tag{20.20}$$

Expansions at $z = 0$ are defined analogously. This can either be done explicitly, as in the following, or via the expansion at ∞ for the Herglotz function $\widetilde{h}(z) := h(-1/z)$. The above remark applies then accordingly.

Definition 20.5. Let h be a Herglotz function. If for $K \geq -1$, there exist real numbers $a_{-1}, a_0, a_1, \ldots, a_K$ such that h can be written as

$$h(z) = \frac{a_{-1}}{z} + a_0 + a_1 z + \cdots + a_K z^K + o(z^K) \qquad \text{as } z \hat{\to} 0 \tag{20.21}$$

then we say that h admits at $z = 0$ an *asymptotic expansion of order K*. □

Remark 20.3. Note that every Herglotz function has asymptotic expansions both at $z = \infty$ and at $z = 0$ of order -1. Indeed, by equalities (20.9) and (20.11), the numbers b_1 and a_1 always exist and are equal to the numbers b from Theorem 20.1 and $-\mu(\{0\})$, respectively.

The following two theorems are the central statements in this context, [14]. We start with the version with positive exponents, which are related to expansions at ∞.

Theorem 20.2. Let h be a Herglotz function. Then, for some integer $N_\infty \geq 0$, the limit

$$\lim_{\varepsilon \to 0^+} \lim_{y \to 0^+} \int_{\varepsilon < |x| < \frac{1}{\varepsilon}} x^{2N_\infty} \operatorname{Im}[h(x + \mathrm{i}y)] \, \mathrm{d}x \tag{20.22}$$

exists as a finite number if and only if the function h admits at $z = \infty$ an asymptotic expansion of order $2N_\infty + 1$. In this case,

$$\lim_{\varepsilon \to 0^+} \lim_{y \to 0^+} \frac{1}{\pi} \int_{\varepsilon < |x| < \frac{1}{\varepsilon}} x^n \operatorname{Im}[h(x + \mathrm{i}y)] \, \mathrm{d}x = \begin{cases} a_{-1} - b_{-1} & n = 0 \\ -b_{-n-1} & 0 < n \leq 2N_\infty \end{cases} \tag{20.23}$$

holds. □

The corresponding result for negative exponents, which are related to the expansions at 0, reads as follows.

Theorem 20.3. Let h be a Herglotz function. Then, for some integer $N_0 \geq 1$, the limit

$$\lim_{\varepsilon \to 0^+} \lim_{y \to 0^+} \int_{\varepsilon < |x| < \frac{1}{\varepsilon}} \frac{\operatorname{Im}[h(x + \mathrm{i}y)]}{x^{2N_0}} \, \mathrm{d}x \tag{20.24}$$

exists (as a finite number) if and only if h admits at $z = 0$ an asymptotic expansion of order $2N_0 - 1$. In this case,

$$\lim_{\varepsilon\to 0^+}\lim_{y\to 0^+}\frac{1}{\pi}\int_{\varepsilon<|x|<\frac{1}{\varepsilon}}\frac{\mathrm{Im}[h(x+\mathrm{i}y)]}{x^p}\,\mathrm{d}x = \begin{cases} a_1 - b_1 & p = 2 \\ a_{p-1} & 2 < p \le 2N_0 \end{cases} \tag{20.25}$$

holds. □

The above integral identities (20.23) and (20.25) are often called *sum rules*.

Remark 20.4. The proof of the above two theorems relies on a version of formula (20.10) for noncontinuous test functions; see, e.g., [34].

Example 20.4. The Herglotz function $h(z) = \tan(z)$ has the asymptotic expansion

$$\tan(z) = z + \frac{z^3}{3} + \frac{2z^5}{15} + \cdots \quad \text{as } z\hat{\to}0 \tag{20.26}$$

and $\tan(z) = \mathrm{i} + o(1)$ as $z\hat{\to}\infty$. We thus find that $a_1 = 1$, $a_3 = 1/3$, $a_5 = 2/15$, and $b_1 = 0$, and the following sum rules

$$\lim_{\varepsilon\to 0^+}\lim_{y\to 0^+}\frac{1}{\pi}\int_{\varepsilon\le|x|\le 1/\varepsilon}\frac{\mathrm{Im}[\tan(x+\mathrm{i}y)]}{x^p}\,\mathrm{d}x = \begin{cases} 1 & p = 2 \\ 1/3 & p = 4 \\ 2/15 & p = 6 \end{cases} \tag{20.27}$$

apply. ■

Remark 20.5. Note that the case of $p = 1$ is not included in Theorem 20.3. In order to guarantee this limit to be finite, it is required that h admits both at ∞ and at the point zero asymptotic expansions of order 1. In this case, the limit equals $a_0 - b_0$.

Remark 20.6. Note that the exponents in (20.22) and (20.24) are even. A corresponding statement for odd exponents, meaning that the existence of the limit is equivalent to the existence of the expansion, does not hold. A counterexample is given in [14, p. 9].

Example 20.5. Note that the assumption that the coefficients in expansions (20.18) and (20.21) are real is essential. Consider, e.g., the function $h(z) = \mathrm{i}$ for $z \in \mathbb{C}^+$, which admits expansions of arbitrary order if non-real coefficients are allowed. However, neither of the limits (20.22) nor (20.24) do exist. Note that this example also shows that not every Herglotz function does admit a sum rule. ■

For symmetric Herglotz functions, we note that the non-zero odd and even ordered coefficients within an asymptotic expansion of a symmetric Herglotz function (20.6) are necessarily real-valued and purely imaginary, respectively. Thus based on Theorems 20.2, 20.3 and Remark 20.5, we stop our expansion at the appearance of the first imaginary term, or the first non-existing term. If the assumptions in both theorems are satisfied, i.e., that both asymptotic series exist, are real-valued to order $2N_0 - 1$ and $2N_\infty + 1$, respectively, we can summarize Theorems 20.2, 20.3 and Remark 20.5 as

$$\frac{2}{\pi}\int_{0^+}^{\infty}\frac{\mathrm{Im}[h(x)]}{x^{2n}}\,\mathrm{d}x := \lim_{\varepsilon\to 0^+}\lim_{y\to 0^+}\frac{2}{\pi}\int_{\varepsilon}^{1/\varepsilon}\frac{\mathrm{Im}[h(x+\mathrm{i}y)]}{x^{2n}}\,\mathrm{d}x = a_{2n-1} - b_{2n-1} \tag{20.28}$$

for $n = -N_\infty, \ldots, N_0$.

20.3 Passive systems

In this section, we show how symmetric Herglotz functions and the corresponding PR functions are related to passive systems.

Physical objects that cannot produce energy are usually considered as passive. However, these objects are not necessarily passive from a system point of view. The crucial point is how the input and the output of the system are defined. For the most part here, we constrain our viewpoint to one-port systems. These systems consist of one input and one output parameter, which can be measured at the *ports* of these systems (see Figure 20.3). Here, the signal enters the system and its response can be measured. A common example of such a system is an electric circuit with two nodes to which we can input a signal, e.g., a current, and measure a voltage. The one-port systems, we regard in this section are linear, continuous, and time-translationally invariant. Time-translation invariance means that the system does not explicitly depend on time; that is, if a system produces the output $v(t)$ from the input $u(t)$ at time t, then a time-shifted input $u(t+\tau)$ gives a shifted output $v(t+\tau)$. These properties characterize systems in convolution form [6].

Definition 20.6. A system with input $u(t)$ and output $v(t)$ in the time-domain is in convolution form if

$$v(t) = (w * u)(t) := \int_{\mathbb{R}} w(\tau)u(t-\tau)\,\mathrm{d}\tau \tag{20.29}$$

where $w(t)$ is the impulse response. □

In this chapter, we restrict ourselves to real-valued systems, i.e., the systems where the impulse response w is real-valued.

There are two ways to define passivity for different types of systems: admittance passivity and scattering passivity [6,35].

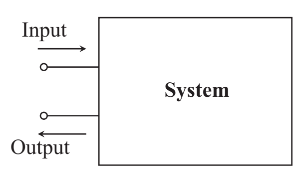


Figure 20.3 A one-port system

Definition 20.7. Consider a convolution system with input signal $u(t)$ and output signal $v(t)$, both of which in general can be complex valued. The system is called *admittance-passive* if

$$\mathscr{W}_{\text{adm}}(T) := \operatorname{Re} \int_{-\infty}^{T} v(t)u(t)^* \, \mathrm{d}t \geq 0 \tag{20.30}$$

for all $T \in \mathbb{R}$ and all $u \in C_0^\infty$ (i.e., smooth functions with compact support). □

Here, $\mathscr{W}_{\text{adm}}(T)$ represents all energy the system has absorbed until the time T. By requiring this quantity to be positive, we say that the system absorbs more energy than it emits, and thus, the system does not produce energy. Passivity also implies that the system is causal [6].

The crucial connection to Herglotz and PR functions is given by the following fact. It can be shown that the impulse response of a passive system has the representation [6]

$$w(t) = b\delta'(t) + \theta(t) \int_{\mathbb{R}} \cos(\xi t) \, \mathrm{d}\mu(\xi) \tag{20.31}$$

where $b \geq 0$, δ' denotes the derivative of the Dirac distribution (cf. [6]), θ is the Heaviside step function (cf. [6]), and the measure μ satisfies the growth condition (20.8). This implies that the Laplace transform of the impulse response (20.31), $W_{\text{adm}}(s)$, is a PR function and that the corresponding symmetric Herglotz function (20.13) has exactly the parameters b and μ.

Let us consider a few examples of passive systems in electromagnetics.

Example 20.6. Input impedance of electrical circuit networks Consider a simple electric one-port circuit containing passive components, meaning that all lumped element components have positive values, e.g., that each resistance R, inductance L, and capacitance C are positive. The input signal to this system is the real-valued electric current $i(t)$ and its output signal is the voltage $v(t)$; see Figure 20.4(a). As an explicit example, consider the simple circuit in Figure 20.4(b). To verify that this system is passive, we evaluate the integral (20.30).

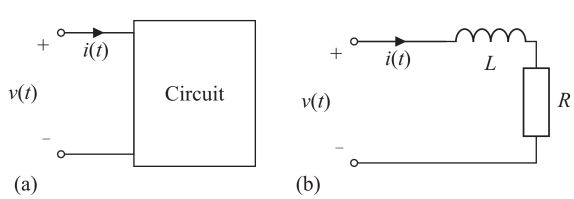


Figure 20.4 Two electrical one-port systems: (a) A general electric circuit. (b) A simple circuit example

Given the input current $i(t)$, the voltage is given by $v(t) = L\frac{\mathrm{d}i(t)}{\mathrm{d}t} + Ri(t) = (w * i)(t)$, where $w = L\delta' + R\delta$ is the impulse response, cf. (20.31). Hence, the integral (20.30) becomes

$$\mathscr{W}_{\mathrm{adm}}(T) = \int_{-\infty}^{T} \left(L\frac{\mathrm{d}i(t)}{\mathrm{d}t} i(t) + Ri(t)^2 \right) \mathrm{d}t = \frac{L}{2} i(T)^2 + R \int_{-\infty}^{T} i(t)^2 \, \mathrm{d}t \geq 0 \tag{20.32}$$

showing that the system is admittance-passive. The transfer function (input impedance), which by definition is the Laplace transform of the impulse response, becomes, in this case, the PR function

$$W_{\mathrm{adm}}(s) := Z_{\mathrm{in}}(s) = sL + R \tag{20.33}$$

The input and output can be interchanged, and by using the voltage as input, the input admittance is given by $Y_{\mathrm{in}}(s) = 1/Z_{\mathrm{in}}(s)$, which is also a PR function. This simple example generalizes to circuit networks composed of arbitrary number and combinations of passive resistors, capacitances, and inductances resulting in rational PR functions [36]. Moreover, it is straightforward to include transformers and transmission lines as well as multiple input and output systems resulting in matrix-valued PR functions [37]. ■

In contrast to this example, in other situations, it is often less clear how to define a system in a suitable way so that it becomes passive. For instance, consider the constitutive relations.

Example 20.7. Constitutive relations Electromagnetic properties of isotropic materials are characterized through constitutive relations with quantities known as permittivity ε and permeability μ, which relate the frequency domain electric $\boldsymbol{E}$ and magnetic $\boldsymbol{H}$ field intensities to the electric $\boldsymbol{D} = \varepsilon\boldsymbol{E}$ and magnetic $\boldsymbol{B} = \mu\boldsymbol{H}$ flux densities, respectively. These constitutive relations are regarded as systems with field intensities and flux densities as input and output signals, respectively. The class of linear, time-translationally invariant, continuous and passive constitutive relations are, hence, related to PR and Herglotz functions. The energy relation in (20.30) is derived by the Poynting's theorem [38,39] and reveals that the pertinent passive systems are the maps from the field intensities to the temporally differentiated flux densities [14,21].

Time-domain admittance passivity implies that in the Laplace domain, the corresponding transfer function is a PR function or, equivalently, a symmetric Herglotz function after some suitable transformations; see Section 20.2. In this domain, time differentiation results in a multiplication with the angular frequency. The constitutive relation $\boldsymbol{D}(s) = \varepsilon(s)\boldsymbol{E}(s)$ is a passive system if $s\varepsilon(s)$ is a PR function. Similarly, permeabilities μ, such that $s\mu(s)$ are PR functions, can be used to model all linear, time-translationally invariant, continuous, and passive models for magnetic media. ■

Let us now consider the second definition of passivity, scattering passivity.

Definition 20.8. Consider a convolution system with input signal $u(t)$ and output signal $v(t)$. The system is called *scattering-passive* if

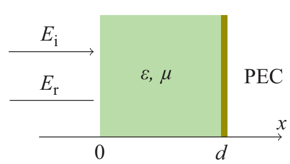
$$\mathscr{W}_{\rm scat}(T) := \int_{-\infty}^{T} (|u(t)|^2 - |v(t)|^2)\, \mathrm{d}t \geq 0 \tag{20.34}$$

for all $T \in \mathbb{R}$ and all $u \in C_0^\infty$. □

Requiring a system to be scattering-passive corresponds to the energy of its output signal always being less than that of its input signal. It can be shown that the transfer function $W_{\rm scat}(s)$ of a scattering-passive system satisfies the relation $|W_{\rm scat}(s)| \leq 1$ for all $s \in \mathbb{C}_+$ [6,14]. Then a suitable rational (Cayley) transformation of the transfer function, namely the function $s \mapsto (1 + W_{\rm scat}(s))/(1 - W_{\rm scat}(s))$, is a PR function.

Let us consider an example of a scattering-passive system.

Example 20.8. Reflection from an isotropic slab placed above a ground plane The reflection coefficient Γ from an isotropic slab characterized by passive permittivity $\varepsilon_{\rm r}$, passive permeability $\mu_{\rm r}$, thickness $d > 0$, and placed above a perfect electric conducting (PEC) plane, see the figure below, can describe a scattering-passive system. Let the incident wave $E_{\rm i}$ be the input signal and the reflected wave $E_{\rm r}$ be the output signal of the system. Assume the reference plane is placed in front of the slab, i.e., $x \leq 0$.



The reflection coefficient for an isotropic slab corresponds to the reflected wave $\boldsymbol{E}_{\rm r} = \Gamma \boldsymbol{E}_{\rm i}$ and is calculated as

$$\Gamma(k) = \frac{\Gamma_0 - \mathrm{e}^{2\mathrm{i}knd}}{1 - \Gamma_0 \mathrm{e}^{2\mathrm{i}knd}} \tag{20.35}$$

where $\Gamma_0 = (\eta - 1)/(\eta + 1)$ is the reflection coefficient at the air–slab interface, k is the wave number, $n = \sqrt{\varepsilon_{\rm r}}\sqrt{\mu_{\rm r}}$ is the refractive index of the slab, and $\eta = \sqrt{\mu_{\rm r}/\varepsilon_{\rm r}}$ is the relative wave impedance of the slab. This function can be shown to map the open right-half of complex plane to the closed unit disk. ■

Admittance-passive and scattering-passive systems are in close correspondence; namely given an admittance-passive system, it is possible to construct a new scattering-passive system, and conversely (see Theorem 5 in [35], as well as [37,40,41]).

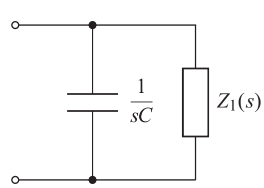
20.4 Sum rules and physical bounds

The integral identities in Theorems 20.2 and 20.3 have a very useful application in the derivation of physical bounds on passive systems; see, e.g., [14]. In the engineering and physics literature, these integral identities appear in various forms and are also often referred to as *sum rules* [9,14]. Typically, the appropriate Herglotz function h is the Fourier transform of a physically realizable real-valued convolution kernel, and which, hence, possesses the symmetry $h(-z^*) = -h(z)^*$, or $p(z^*) = p(z)^*$ for a PR function. For Herglotz functions, the integral identities are given on the real axis where $z = x$ is commonly interpreted as angular frequency ω (in rad/s), wave number $k = \omega/c_0$ (in m^{-1}), or as wavelength $\lambda = 2\pi/k$ (in m). For PR functions, the imaginary axis $z = \mathrm{i}y$ is instead playing this role.

In many practical electromagnetic applications, it is reasonable to assume (or even to acquire by measurements) some partial knowledge regarding the low- and/or high-frequency asymptotic expansions of the corresponding Herglotz function, such as the static and the optical responses of a material, or a structure. In these cases, the sum rules and the corresponding integral identities can be used to obtain inequalities by constraining the integration interval to a finite bandwidth in the frequency (or wavelength) domain, and thereby yielding useful physical limitations in a variety of applications. Typical examples are with matching networks [15], radar absorbers [16], high-impedance surfaces [20], passive metamaterials [21], scattering [17,42], antennas [18,22,23,43], reflection coefficients [44], waveguides [45], periodic structures [46], etc.

Three examples are given next to illustrate typical situations where sum rules can be used to derive physical limitations on passive electromagnetic systems. In special situations, such results have previously been derived using residue calculus and Kramers–Kronig relations [16,36,38,47–49], whereas here, we consistently use the general result (20.28).

Example 20.9. The resistance-integral theorem Consider a passive circuit consisting of a parallel connection of a capacitance C and an impedance $Z_1(s)$ that does not contain a shunt capacitance (i.e., $Z_1(0)$ is finite and $Z_1(s)$ does not have a zero at $s = \infty$); see the figure below. Then the input impedance of this circuit is given by $Z(s) = 1/(sC + 1/Z_1(s))$, which is a PR function in the Laplace variable $s \in \mathbb{C}_+$, and hence the system is admittance passive.



The asymptotic expansions are $Z(s) = Z_1(0) + o(s)$ as $s\hat{\to}0$ and $Z(s) = 1/(sC) + o(s^{-1})$ as $s\hat{\to}\infty$. Here, the corresponding Herglotz function is $h(\omega) := \mathrm{i}Z(-\mathrm{i}\omega)$ where $s = -\mathrm{i}\omega$ and $\omega \in \mathbb{C}^+$. Its low- and high-frequency asymptotics are

$$h(\omega) = o(\omega^{-1}) \quad \text{as } \omega\hat{\to}0 \quad \text{and} \quad h(\omega) = -\frac{1}{\omega C} + o(\omega^{-1}) \quad \text{as } \omega\hat{\to}\infty \quad (20.36)$$

In terms of (20.21) and (20.18), we have $a_{-1} = 0$ and $b_{-1} = -1/C$, and thus the sum rule (20.28) with $n = 0$ gives

$$\frac{2}{\pi}\int_{0^+}^{\infty} \mathrm{Re}[Z(-\mathrm{i}\omega)]\,\mathrm{d}\omega = \frac{2}{\pi}\int_{0^+}^{\infty} \mathrm{Im}[h(\omega)]\,\mathrm{d}\omega = a_{-1} - b_{-1} = \frac{1}{C} \qquad (20.37)$$

By integrating only over a finite frequency interval $\Omega := [\omega_1, \omega_2]$, and estimating this integral from the following, we obtain the bound

$$\Delta\omega \inf_{\omega\in\Omega} \mathrm{Re}[Z(-\mathrm{i}\omega)] \leq \int_{0^+}^{\infty} \mathrm{Re}[Z(-\mathrm{i}\omega)]\,\mathrm{d}\omega = \frac{\pi}{2C} \qquad (20.38)$$

where $\Delta\omega := \omega_2 - \omega_1$. Consequently, inequality (20.38) limits the product between the bandwidth and the minimum resistance over the given frequency interval; see also [47]. ■

Compositions of Herglotz functions can be used to construct new Herglotz functions and, hence, also new sum rules. Here, we illustrate this for a case where the minimal temporal dispersion for metamaterials is determined. The problem is first transformed to the problem of determining the minimum amplitude of a Herglotz function over a bandwidth [14,21].

Example 20.10. Metamaterials and temporal dispersion When a dielectric medium is specified to have inductive properties (i.e., has negative permittivity) over a given bandwidth, it is regarded as a metamaterial. A given negative permittivity value at a single frequency is always possible to achieve. For instance, the plasmonic resonances in small metal particles can be readily explained by using common Drude or Lorentz models, etc. However, when a constant negative permittivity value is prescribed over a given bandwidth, the passivity of the material will imply severe bandwidth limitations; see, e.g., [21].

To derive these limitations based on the general theory of Herglotz functions, we start by considering the following general situation: Let $F(x) := -h_0(x)$ be the negative of a fixed Herglotz function h_0 that can be extended continuously to a neighborhood of the compact interval $\Omega \subset \mathbb{R}$ and has the large argument asymptotics $h_0(z) = b_1^0 z + o(z)$ as $z\hat{\to}\infty$. Let h denote an arbitrary Herglotz function with the same continuity property on the real line and satisfying the asymptotics $h(z) = b_1 z + o(z)$ as $z\hat{\to}\infty$. We aim to derive a lower bound for the error norm

$$\|h - F\|_{\mathrm{L}^\infty(\Omega)} := \sup_{x\in\Omega} |h(x) - F(x)| \qquad (20.39)$$

To this end, an auxiliary Herglotz function $h_\Delta(z)$, cf. [21], is defined by

$$h_\Delta(z) := \frac{1}{\pi}\int_{-\Delta}^{\Delta}\frac{1}{\xi - z}\,\mathrm{d}\xi = \frac{1}{\pi}\mathrm{Log}\frac{z-\Delta}{z+\Delta} = \begin{cases} \mathrm{i} + o(1) & \text{as } z\hat{\to}0 \\ \dfrac{-2\Delta}{\pi z} + o(z^{-1}) & \text{as } z\hat{\to}\infty \end{cases} \tag{20.40}$$

Note that $\mathrm{Im}[h_\Delta(z)] \geq \frac{1}{2}$ for $|z| \leq \Delta$ and $\mathrm{Im}[z] \geq 0$. Next, consider the composite Herglotz function $h_1(z) := h_\Delta(h(z) + h_0(z))$, where $h(z) + h_0(z) = (b_1 + b_1^0)z + o(z)$ as $z\hat{\to}\infty$ yields the asymptotics

$$h_1(z) = o(z^{-1}) \text{ as } z\hat{\to}0 \quad \text{and} \quad h_1(z) = \frac{-2\Delta}{\pi(b_1 + b_1^0)}z^{-1} + o(z^{-1}) \text{ as } z\hat{\to}\infty \tag{20.41}$$

The sum rule (20.28) with $n = 0$ is given by

$$\frac{2}{\pi}\int_{0+}^{\infty}\mathrm{Im}[h_1(x)]\,\mathrm{d}x = a_{-1} - b_{-1} = \frac{2\Delta}{\pi(b_1 + b_1^0)} \tag{20.42}$$

Choosing $\Delta := \sup_{x\in\Omega}|h(x) + h_0(x)|$, the following integral inequalities follow

$$\frac{1}{\pi}|\Omega| \leq \frac{2}{\pi}\int_\Omega \underbrace{\mathrm{Im}[h_1(x)]}_{\geq\frac{1}{2}}\,\mathrm{d}x \leq \frac{2}{\pi}\int_{0+}^{\infty}\mathrm{Im}[h_1(x)]\,\mathrm{d}x = \frac{2\sup_{x\in\Omega}|h(x) + h_0(x)|}{\pi(b_1 + b_1^0)} \tag{20.43}$$

or

$$\|h + h_0\|_{\mathrm{L}^\infty(\Omega)} \geq (b_1 + b_1^0)\frac{1}{2}|\Omega|, \quad \text{where } |\Omega| = \int_\Omega \mathrm{d}x \tag{20.44}$$

Consider now a dielectric metamaterial with a constant, real-valued, and negative target permittivity $\varepsilon_\mathrm{t} < 0$ to be approximated over an interval Ω. In this case, $F(z) = z\varepsilon_\mathrm{t}$ and $h_0(z) = -F(z)$ with $b_1^0 = -\varepsilon_\mathrm{t}$. Let $\varepsilon(z)$ be the permittivity function of the approximating passive dielectric material, and $h(z) = z\varepsilon(z)$ the corresponding Herglotz function with $b_1 = \varepsilon_\infty$, the assumed high-frequency permittivity of the material, and the approximation interval $\Omega = \omega_0[1 - B/2, 1 + B/2]$, where ω_0 is the center frequency and B the relative bandwidth with $0 < B < 2$. The resulting physical bound obtained from (20.44) is given by

$$\|\varepsilon(\,\cdot\,) - \varepsilon_\mathrm{t}\|_{\mathrm{L}^\infty(\Omega)} \geq \frac{(\varepsilon_\infty - \varepsilon_\mathrm{t})B}{2 + B} \tag{20.45}$$

see also [21]. Note that the variable x corresponds here to angular frequency, also commonly denoted as ω (in rad/s). ■

Scattering passive systems have transfer functions that map $\mathbb{C}^+$ to the unit disk. To use (20.28), we first construct a Herglotz (or PR) function by mapping the unit disk to $\mathbb{C}^+$. This map can be made in many different ways and the particular choice depends on the asymptotic expansion and the physical interpretation of the system.

The Cayley transform, logarithm, and addition are most common in applications. Here, the reflection coefficient of a radar absorber is considered. It is desired to bound the magnitude of the reflection coefficient and, hence, the logarithm is used to construct a Herglotz function.

Example 20.11. Sum rules for passive radar absorbers The reflection coefficient $\Gamma(k)$ for an isotropic slab with permeability $\mu(k)$ and permittivity $\varepsilon(k)$ was derived in Example 20.8, where $k = \omega/c_0$ is the wave number. Assuming that $\mu(k) = \mu_\mathrm{s} + O(k)$ and $\varepsilon(k) = \varepsilon_\mathrm{s} + O(k)$, where μ_s and ε_s are the corresponding static values, it can be shown that the reflection coefficient has the following asymptotic expansions

$$\Gamma(k) = -1 - \mathrm{i}2kd\mu_\mathrm{s} + O(k^2) \text{ as } k\hat{\to}0 \quad \text{and} \quad \Gamma(k) = O(1) \text{ as } k\hat{\to}\infty \tag{20.46}$$

The function $\Gamma(k)$ is the Fourier transform of a scattering passive convolution kernel, so $\Gamma(k)$ is an analytic function with $|\Gamma(k)| \leq 1$ for $k \in \mathbb{C}^+$. Hence it has a representation of the form $\Gamma(k) = -B(k)\mathrm{e}^{\mathrm{i}h(k)}$, where $B(k)$ is a Blaschke product (cf. [50]) and $h(k)$ a Herglotz function [14]. With an appropriately chosen branch of the logarithm, h can thus be written

$$h(k) := -\mathrm{i}\,\log\left(-\Gamma(k)\prod_n \frac{1 - k/k_n^*}{1 - k/k_n}\right) \tag{20.47}$$

where k_n are the zeros of $\Gamma(k)$ with $\mathrm{Im}[k_n] > 0$. The Blaschke product is used to remove these zeros from the upper half plane and the negative sign is chosen here to make the Herglotz function $h(k)$ symmetric.

The asymptotic expansions of $h(k)$ are given by

$$h(k) = k\left(2d\mu_\mathrm{s} + 2\sum_n \mathrm{Im}\left[\frac{1}{k_n}\right]\right) + o(k) \quad \text{as } k\hat{\to}0 \tag{20.48}$$

and $h(k) = o(k)$ as $k\hat{\to}\infty$. Therefore, there is a sum rule (20.28) yielding

$$\frac{2}{\pi}\int_{0^+}^{\infty} \frac{\mathrm{Im}[h(k)]}{k^2}\,\mathrm{d}k = 2d\mu_\mathrm{s} + 2\sum_n \mathrm{Im}\left[\frac{1}{k_n}\right] \tag{20.49}$$

Since $\mathrm{Im}[1/k_n] < 0$, the following inequality is obtained:

$$\frac{2}{\pi}\int_{0^+}^{\infty} \frac{-\ln|\Gamma(k)|}{k^2}\,\mathrm{d}k \leq 2d\mu_\mathrm{s}, \quad \text{or} \quad \int_{0^+}^{\infty} -\ln|\Gamma(\lambda)|\,\mathrm{d}\lambda \leq 2\pi^2 d\mu_\mathrm{s} \tag{20.50}$$

where we used $\lambda = 2\pi/k$ to arrive at the final formulation, which was originally given in [16]. Note that the inequality in (20.50) gives also the following bound on the absorption parameter $1/|\Gamma(\lambda)|$ over the interval $[\lambda_1, \lambda_2]$

$$\Delta\lambda \inf_{\lambda\in[\lambda_1,\lambda_2]} \ln\frac{1}{|\Gamma(\lambda)|} \leq \int_{0^+}^{\infty} \ln\frac{1}{|\Gamma(\lambda)|}\,\mathrm{d}\lambda \leq 2\pi^2 d\mu_\mathrm{s} \tag{20.51}$$

where $\Delta\lambda = \lambda_2 - \lambda_1$.

The derivation above is, for simplicity, presented for homogeneous slabs similar to the multilayer case originally presented in [16]. The sum rule and physical bounds are generalized to arbitrary inhomogeneous periodic structures using the low-frequency expansion in [20] together with passivity and expansion of the reflected field in Floquet modes. This derivation shows that (20.51) is valid for the general case. An extension of the result (20.51) has applications in the evaluation of bandwidth performance of array antennas [22,23]. ■

20.5 Convex optimization and physical bounds

Convex optimization [27,51] based on the Herglotz function representation (20.7) can be used to approximate and identify passive systems. However, to facilitate the computation of a numerical solution using a software such as CVX [51], it is necessary to first impose some a-priori constraints on the class of approximating Herglotz functions. In particular, we are interested here in Herglotz functions that are known to be locally Hölder continuous on some given intervals on the real line. Hence, a passive approximation problem is considered where the target function F is an arbitrary complex-valued continuous function defined on an approximation domain $\Omega \subset \mathbb{R}$ consisting of a finite union of closed and bounded intervals of the real axis. The norms used, denoted by $\|\cdot\|_{\mathrm{L}^p(w,\Omega)}$, are weighted $\mathrm{L}^p(\Omega)$-norms [50] which are defined here by using a positive continuous weight function w on Ω, and where $1 \leq p \leq \infty$.

Here, the approximating function h is the Hölder continuous extension (to Ω) of some Herglotz function generated by an absolutely continuous measure μ having a density μ' which is Hölder continuous on the closure $\overline{U}$ of an arbitrary neighborhood $U \supset \Omega$ of the approximation domain. The function μ' is Hölder continuous with Hölder exponent α, meaning that $|\mu'(\tau) - \mu'(\varsigma)| \leq C\,|\tau - \varsigma|^{\alpha}$ for all $\tau, \varsigma \in \overline{U}$, where $0 < \alpha < 1$ is fixed, and $C > 0$ is an arbitrary constant. The corresponding Hölder space is denoted $\mathrm{C}^{0,\alpha}(\overline{U})$; see, e.g., [52, pp. 94–104]. It can be shown that the Hilbert transform integral operator $\mathscr{H}$, defined similarly as in (20.16), is a bounded operator $\mathscr{H}\colon \mathrm{C}^{0,\alpha}(\overline{U}) \to \mathrm{C}^{0,\alpha}(\Omega)$; cf., e.g., Theorem 7.6 and Corollary 7.7 on pp. 101–102 in [52] and see also [9]. By assumption, both the real and the imaginary parts of h are continuous functions on Ω. Moreover, it holds that $\mathrm{Im}[h] = \pi\mu'$ on $\overline{U}$ (cf. [13, p. 7]) and, due to the Hölder continuity of the density μ' on $\overline{U}$, the real part is given by the associated Hilbert transform [52], similar as in (20.17). Now, the impulse response of a passive system can usually be considered as real-valued. In this case, the approximating Herglotz function h can be assumed to be symmetric and its real part admits the representation

$$\mathrm{Re}[h(x)] = bx + \mathrm{p.v.}\int_{\mathbb{R}} \frac{\mu'(\tau)}{\tau - x}\,\mathrm{d}\tau \quad \text{for } x \in \Omega \tag{20.52}$$

where $\mathrm{Im}[h(x)] = \pi\mu'(x)$ for $x \in \Omega$.

The continuity of h on Ω implies that the norm $\|h\|_{L^p(w,\Omega)}$ is well-defined for $1 \le p \le \infty$. An approximation problem of interest can now be formulated in terms of the greatest lower bound on the approximation error, defined by

$$d := \inf_h \|h - F\|_{L^p(w,\Omega)} \tag{20.53}$$

where the infinum is taken over all Herglotz functions h generated by a measure having a Hölder continuous density on $\overline{U}$.

In general, a best approximation achieving the bound d in (20.53) does not exist. In practice, however, the problem is approached by using numerical algorithms such as CVX [51], solving finite-dimensional approximation problems using B-splines, with the number of basis functions N fixed during the optimization, cf. [24,25]. Here, a B-spline of order $m \ge 2$ is an $m-2$ times continuously differentiable and compactly supported positive basis spline function consisting of piecewise polynomial functions of order $m-1$, i.e., linear, quadratic, cubic, etc., and which is defined by $m+1$ break points [53]. Let us now consider a discretization of the problem expressed in (20.53), which is based on an arbitrary, finite partition of the approximation domain Ω. Let

$$\pi\mu'(x) = \sum_{n=1}^{N} \zeta_n \left(p_n(x) + p_n(-x)\right) \tag{20.54}$$

for $x \in \mathbb{R}$ be a finite B-spline expansion of $\mathrm{Im}\,[h(x)]$, where ζ_n are optimization variables for $n = 1, \ldots, N$, and $p_n(x)$ are B-spline basis functions of fixed order m which are defined on the given partition. The real part $\mathrm{Re}[h(x)]$ for $x \in \Omega$ is then given by (20.52), and can be expressed as

$$\mathrm{Re}[h(x)] = bx - \frac{\zeta_0}{x} + \sum_{n=1}^{N} \zeta_n \left(\hat{p}_n(x) - \hat{p}_n(-x)\right), \quad x \in \Omega \tag{20.55}$$

where $\hat{p}_n(x)$ is the (negative) Hilbert transform of the B-spline function $p_n(x)$ and where a point mass at $x = 0$ with amplitude c_0 has been included; see (20.13). Any other a-priori assumed point masses can be included in a similar way. Explicit formulas for general B-splines and their Hilbert transforms are given in [24]. As an example, a piecewise linear ("roof-top") B-spline on a uniform partition is given by $p_n(x) = p(x - n\Delta)$, where

$$p(x) = \begin{cases} 1 - |x|/\Delta & |x| \le \Delta \\ 0 & |x| > \Delta \end{cases} \tag{20.56}$$

for $x \in \mathbb{R}$ and $\Delta > 0$, and its (negative) Hilbert transform $\hat{p}_n(x) = \hat{p}(x - n\Delta)$, where

$$\hat{p}(x) = \frac{1}{\pi\Delta}\left(2x \ln|x| - (x-\Delta)\ln|x-\Delta| - (x+\Delta)\ln|x+\Delta|\right) \tag{20.57}$$

for $x \in \mathbb{R}$. Note that the logarithmic singularities in (20.57) cancel, and the function $\hat{p}(x)$ is a continuous function on $\mathbb{R}$.

Consider now the following convex optimization problem:

$$\begin{array}{ll} \text{minimize} & \|h - F\|_{\mathrm{L}^p(w,\Omega)} \\ \text{subject to} & \zeta_n \geq 0, \ \text{for } n = 0, \ldots, N \\ & b \geq 0, \end{array} \tag{20.58}$$

where the optimization is over the variables $(\zeta_0, \zeta_1, \ldots, \zeta_N, b)$. Note that the objective function in (20.58) above is the norm of an affine form in the optimization variables. Hence, the objective function is a convex function in the variables $(\zeta_0, \zeta_1, \ldots, \zeta_N, b)$.

The uniform continuity of all functions involved implies that the solution to (20.58) can be approximated within an arbitrary accuracy by discretizing the approximation domain Ω (and the computation of the norm) using only a finite number of sample points. The corresponding numerical problem (20.58) can now be solved efficiently by using the CVX MATLAB® software for disciplined convex programming [51]. The convex optimization formulation (20.58) offers a great advantage in the flexibility in which additional or alternative convex constraints and formulations can be implemented; see also [24,25].

Example 20.12. A canonical example of convex optimization is with the passive approximation of metamaterials; see also [21,24,25]. As in Example 20.10, the variable x corresponds here to angular frequency, also commonly denoted as ω (in rad/s). A typical application is with the study of optimal plasmonic resonances in small structures (or particles) for which the absorption cross section can be approximated by

$$\sigma_{\text{abs}} \approx k\text{Im}[\gamma] \tag{20.59}$$

where $k = 2\pi/\lambda$ is the wave number of free space, λ the wavelength, and γ is the electric polarizability of the particle; see [54]. As, e.g., the polarizability of a dielectric sphere with radius a is given by $\gamma(x) = 4\pi a^3(\varepsilon(x) - 1)/(\varepsilon(x) + 2)$, where $\varepsilon(x)$ is the permittivity function of the dielectric material inside the sphere.

A surface plasmon resonance is obtained when $\varepsilon(x) \approx -2$, and, hence, we specify that the target permittivity of our metamaterial is $\varepsilon_{\text{t}} = -2$. However, a metamaterial with a negative real part cannot, in general, be implemented as a passive material over a given bandwidth, cf. [21]. Based on the theory of Herglotz functions and associated sum rules, the physical bound in (20.45) can be derived, where ε_∞ is the high-frequency permittivity of the material, $\varepsilon_{\text{t}} < \varepsilon_\infty$, $\Omega = \omega_0[1 - B/2, 1 + B/2]$, ω_0 the center frequency, and B the relative bandwidth with $0 < B < 2$, cf. [21]. The convex optimization formulation (20.58) can be used to study passive realizations (20.54) and (20.55) that satisfies the bound (20.45) as close as possible. Here, the approximating Herglotz function is $h(x) = x\varepsilon(x)$, the target function $F(x) = x\varepsilon_{\text{t}}$, ζ_0 the amplitude of a point mass at $x = 0$, $b = \varepsilon_\infty$, and a weighted norm is used defined by $\|f\|_{\mathrm{L}^\infty(w,\Omega)} = \max_{x\in\Omega} |f(x)/x|$ assuming that $0 \notin \Omega$.

In Figure 20.5 is shown the absorption cross section $\sigma_{\text{abs}}^{\text{D}}$ of a small ($a = 5\,\text{nm}$) gold nanosphere modeled by using a dielectric Drude model with static conductivity $\sigma_0 = 4.52 \cdot 10^7\,\text{S/m}$ and collision frequency $\nu = 4.449 \cdot 10^{12}\,\text{Hz}$; see [55]. The plot also shows the maximal bandwidth (MB) for the corresponding Drude peak resonance based on the associated sum rule $\Delta\lambda = 4\pi^3 a^3/\max \sigma_{\text{abs}}^{\text{D}}$ corresponding to a relative

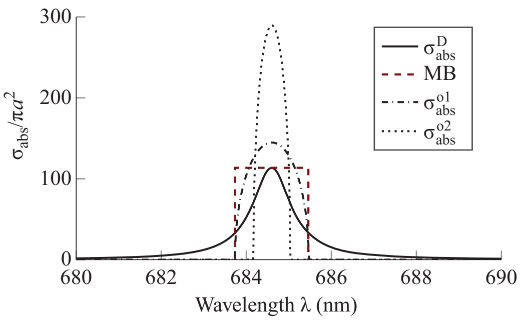


Figure 20.5 *Absorption cross section* $\sigma^{\mathrm{D}}_{\mathrm{abs}}$ *of a small (a* = 5 nm*) gold nanosphere modeled by using a dielectric Drude model. The red dashed line (MB) gives the maximal bandwidth for the corresponding peak resonance based on the associated sum rule* $\Delta\lambda = 4\pi^3 a^3 / \max \sigma^{\mathrm{D}}_{\mathrm{abs}}$. *Here,* $\sigma^{\mathrm{o1}}_{\mathrm{abs}}$ *and* $\sigma^{\mathrm{o2}}_{\mathrm{abs}}$ *correspond to optimized permittivity functions for relative bandwidths* $B = 2.5 \cdot 10^{-3}$ *and* $B = 1.25 \cdot 10^{-3}$*, respectively*

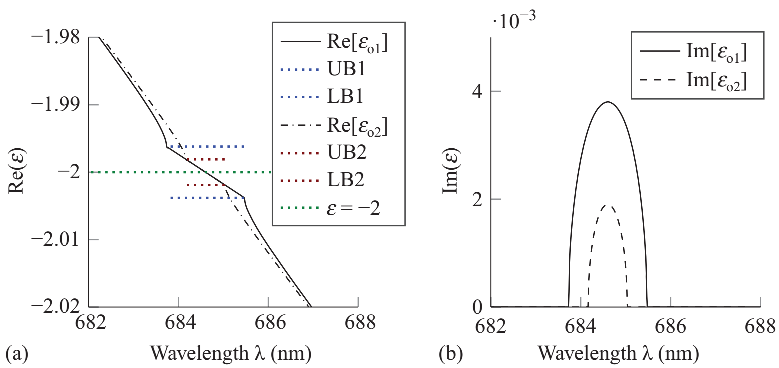


Figure 20.6 *Real and imaginary parts of the optimized permittivity functions approximating a metamaterial with* $\varepsilon_{\mathrm{t}} = -2$. *Here,* $\varepsilon_{\mathrm{o1}}$ *and* $\varepsilon_{\mathrm{o2}}$ *are optimized using* (20.58) *and linear B-splines, for the relative bandwidths* $B = 2.5 \cdot 10^{-3}$ *and* $B = 1.25 \cdot 10^{-3}$*, respectively. UB1-2 and LB1-2 denote the corresponding upper and lower physical bounds based on* (20.45)

bandwidth of $B = 2.5 \cdot 10^{-3}$. Finally, the plot shows the absorption cross sections $\sigma^{\mathrm{o1}}_{\mathrm{abs}}$ and $\sigma^{\mathrm{o2}}_{\mathrm{abs}}$ in two examples, where the permittivity functions have been optimized using (20.58) and linear B-splines for the relative bandwidths $B = 2.5 \cdot 10^{-3}$ and $B = 1.25 \cdot 10^{-3}$, respectively. The resulting optimal permittivity functions are plotted in Figure 20.6 together with the fundamental bound (20.45). Note that there is also a point mass with amplitude ζ_0 at $x = 0$ which is not seen in this plot. All plots are made with respect to wavelength $\lambda = 2\pi/k$. ■

20.6 Conclusions

We have illustrated how the theory of Herglotz functions provides essential information about electromagnetic systems and engineering challenges, both by using this theory on well-known problems and recent results. Sum rules and convex optimization utilize the inherent constraints due to passivity, linearity, and time-translational invariance of the electromagnetic applications to obtain fundamental physical bounds. Furthermore, these techniques not only give physical bounds, but in several cases have been shown to be predictive as a tool in the design of electromagnetic structures.

Acknowledgments

This work was supported by the Swedish Foundation for Strategic Research (SSF) under the program Applied Mathematics and the project Complex Analysis and Convex Optimization for EM Design.

References

[1] N.I. Akhiezer and I.M. Glazman. *Theory of Linear Operators in Hilbert Space*. Vol. 1. Frederick Ungar Publishing Co., 1961.

[2] M. Reed and B. Simon. *Methods of Modern Mathematical Physics . Vol. I: Functional Analysis*. Academic Press, 1972.

[3] N.I. Akhiezer. *The Classical Moment Problem*. Oliver and Boyd, 1965.

[4] B. Simon. "The classical moment problem as a self-adjoint finite difference operator." *Advances in Mathematics* 137.1 (1998), pp. 82–203.

[5] J.C. Willems. "Dissipative dynamical systems part II: linear systems with quadratic supply rates." *Arch. Rational Mech. Anal.* 45.5 (1972), pp. 352–393.

[6] A.H. Zemanian. *Distribution Theory and Transform Analysis: An Introduction to Generalized Functions, with Applications*. McGraw-Hill, 1965.

[7] O. Brune. "Synthesis of a finite two-terminal network whose driving-point impedance is a prescribed function of frequency." *MIT J. Math. Phys.* 10 (1931), pp. 191–236.

[8] S. Darlington. "Synthesis of reactance 4-poles which produce prescribed insertion loss characteristics." *Journal of Mathematics and Physics* 18 (1939), pp. 275–353.

[9] F.W. King. *Hilbert Transforms Vol. I–II*. Cambridge University Press, 2009.

[10] H.M. Nussenzveig. *Causality and Dispersion Relations*. Academic Press, 1972.

[11] G.W. Milton, M. Cassier, O. Mattei, M. Milgrom, and A. Welters. *Extending the Theory of Composites to Other Areas of Science*. Milton-Patton Publishers, 2016.
[12] G.W. Milton. *The Theory of Composites*. Cambridge University Press, 2002.
[13] I.S. Kac and M.G. Krein. "R-functions—analytic functions mapping the upper halfplane into itself." *Amer. Math. Soc. Transl. (2)* 103 (1974), pp. 1–18.
[14] A. Bernland, A. Luger, and M. Gustafsson. "Sum rules and constraints on passive systems." *J. Phys. A: Math. Theor.* 44.14 (2011), p. 145205.
[15] R.M. Fano. "Theoretical limitations on the broadband matching of arbitrary impedances." *Journal of the Franklin Institute* 249.1,2 (1950), pp. 57–83 and 139–154.
[16] K.N. Rozanov. "Ultimate thickness to bandwidth ratio of radar absorbers." *IEEE Trans. Antennas Propag.* 48.8 (2000), pp. 1230–1234.
[17] C. Sohl, M. Gustafsson, and G. Kristensson. "Physical limitations on broadband scattering by heterogeneous obstacles." *J. Phys. A: Math. Theor.* 40 (2007), pp. 11165–11182.
[18] M. Gustafsson, C. Sohl, and G. Kristensson. "Physical limitations on antennas of arbitrary shape." *Proc. R. Soc. A* 463 (2007), pp. 2589–2607.
[19] C.R. Brewitt-Taylor. "Limitation on the bandwidth of artificial perfect magnetic conductor surfaces." *Microwaves, Antennas & Propagation, IET* 1.1 (2007), pp. 255–260.
[20] M. Gustafsson and D. Sjöberg. "Physical bounds and sum rules for high-impedance surfaces." *IEEE Trans. Antennas Propag.* 59.6 (2011), pp. 2196–2204.
[21] M. Gustafsson and D. Sjöberg. "Sum rules and physical bounds on passive metamaterials." *New Journal of Physics* 12 (2010), p. 043046.
[22] J.P. Doane, K. Sertel, and J.L. Volakis. "Matching bandwidth limits for arrays backed by a conducting ground plane." *IEEE Trans. Antennas Propag.* 61.5 (2013), pp. 2511–2518.
[23] B.L.G. Jonsson, C.I. Kolitsidas, and N. Hussain. "Array antenna limitations." *Antennas and Wireless Propagation Letters, IEEE* 12 (2013), pp. 1539–1542.
[24] Y. Ivanenko, M. Gustafsson, B.L.G. Jonsson, *et al. Passive Approximation and Optimization with B-splines*. Tech. rep. URN: urn:nbn:se:lnu:diva-63878. Linnaeus University, 2017.
[25] S. Nordebo, M. Gustafsson, B. Nilsson, and D. Sjöberg. "Optimal realizations of passive structures." *IEEE Trans. Antennas Propag.* 62.9 (2014), pp. 4686–4694.
[26] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. SIAM, 1994.
[27] S.P. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge Univ. Pr., 2004.
[28] L.V. Ahlfors. *Complex Analysis*. McGraw-Hill Book Company, 1979.
[29] F. Gesztesy and E. Tsekanovskii. "On matrix-valued Herglotz functions." *Mathematische Nachrichten* 218.1 (2000), pp. 61–138.

[30] M. Abramowitz and I.A. Stegun, eds. *Handbook of Mathematical Functions*. Applied Mathematics Series No. 55. National Bureau of Standards, 1970.

[31] N.I. Akhiezer and I.M. Glazman. *Theory of Linear Operators in Hilbert Space*. Vol. 2. Frederick Ungar Publishing Co., 1963.

[32] G.E. Shilov and B.L. Gurevich. *Integral, Measure and Derivative: A Unified Approach*. *Dover Books on Mathematics*. Dover Publications, 2013.

[33] J.B. Garnett. *Bounded Analytic Functions*. Revised first. Springer-Verlag, 2007.

[34] P. Henrici. *Applied and Computational Complex Analysis*, *Volume 3: Discrete Fourier Analysis, Cauchy Integrals, Construction of Conformal Maps, Univalent Functions*. John Wiley & Sons, 1993.

[35] M. Wohlers and E. Beltrami. "Distribution theory as the basis of generalized passive-network analysis." *IEEE Transactions on Circuit Theory* 12.2 (1965), pp. 164–170.

[36] E.A. Guillemin. *Synthesis of Passive Networks*. John Wiley & Sons, 1957.

[37] D. Youla, L. Castriota, and H. Carlin. "Bounded real scattering matrices and the foundations of linear passive network theory." *IRE Transactions on Circuit Theory* 6.1 (1959), pp. 102–124.

[38] J.D. Jackson. *Classical Electrodynamics*. Third Edition. John Wiley & Sons, 1999.

[39] G. Kristensson. *Scattering of Electromagnetic Waves by Obstacles*. SciTech Publishing, an Imprint of the IET, 2016.

[40] V.S. Vladimirov. *Equations of Mathematical Physics*. Mir Publishers, 1984.

[41] A.H. Zemanian. "An *n*-port realizability theory based on the theory of distributions." *IEEE Transactions on Circuit Theory* 10.2 (1963), pp. 265–274.

[42] A. Bernland, M. Gustafsson, and S. Nordebo. "Physical limitations on the scattering of electromagnetic vector spherical waves.' *J. Phys. A: Math. Theor.* 44.14 (2011), p. 145401.

[43] C. Sohl and M. Gustafsson. "A priori estimates on the partial realized gain of ultra-wideband (UWB) antennas." *Quart. J. Mech. Appl. Math.* 61.3 (2008), pp. 415–430.

[44] M. Gustafsson. "Sum rules for lossless antennas." *IET Microwaves, Antennas & Propagation* 4.4 (2010), pp. 501–511.

[45] I. Vakili, M. Gustafsson, D. Sjoberg, R. Seviour, M. Nilsson, and S. Nordebo. "Sum rules for parallel-plate waveguides: experimental results and theory." *IEEE Trans. Microwave Theory Tech.* 62.11 (2014), pp. 2574–2582.

[46] M. Gustafsson, I. Vakili, S.E.B. Keskin, D. Sjöberg, and C. Larsson. "Optical theorem and forward scattering sum rule for periodic structures." *IEEE Trans. Antennas Propag.* 60.8 (2012), pp. 3818–3826.

[47] H.W. Bode. *Network Analysis and Feedback Amplifier Design*, 1945. Van Nostrand, 1945.

[48] L.D. Landau, E.M. Lifshitz, and L.P. Pitaevskiĭ. *Electrodynamics of Continuous Media*. Second Edition. Pergamon Press, 1984.

[49] O. Wing. *Classical Circuit Theory*. Springer, 2008.

[50] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, 1987.
[51] M. Grant and S. Boyd. *CVX: A System for Disciplined Convex Programming*, Release 2.0, © 2012 CVX Research, Inc., Austin, TX.
[52] R. Kress. *Linear Integral Equations*. Second Edition. Springer-Verlag, 1999.
[53] C. De Boor. "On calculating with B-splines." *Journal of Approximation Theory* 6.1 (1972), pp. 50–62.
[54] C.F. Bohren and D.R. Huffman. *Absorption and Scattering of Light by Small Particles*. John Wiley & Sons, 1983.
[55] M.G. Blaber, M.D. Arnold, and M.J. Ford. "Search for the ideal plasmonic nanoshell: the effects of surface scattering and alternatives to gold and silver." *J. Phys. Chem. C* 113.8 (2009), pp. 3041–3045.