A Note on BIBO Stability

Michael Unser

***Abstract*—The statements on the BIBO stability of continuous-time convolution systems found in engineering textbooks are often either too vague (because of lack of hypotheses) or mathematically incorrect. What is more troubling is that they usually exclude the identity operator. The purpose of this note is to clarify the issue while presenting some fixes. In particular, we show that a linear shift-invariant system is BIBO-stable in the L_∞-sense if and only if its impulse response is included in the space of bounded Radon measures, which is a superset of $L_1(\mathbb{R})$ (Lebesgue's space of absolutely integrable functions). As we restrict the scope of this characterization to the convolution operators whose impulse response is a measurable function, we recover the classical statement.**

I. Introduction

A statement that is made in most courses on the theory of linear systems as well as in the English version of Wikipedia[1] is that a convolution operator is stable in the BIBO sense (bounded input and bounded output) if and only if its impulse response is absolutely summable/integrable. While the proof of this equivalence is fairly straightforward for discrete-time systems, there seems to be some confusion in the continuous domain (see Appendix B for specific references), especially since the above statement excludes the identity operator, whose impulse response is the Dirac distribution δ. Since δ is not a measurable function in the sense of Lebesgue (see explanations in Appendix A) and hence not a member of $L_1(\mathbb{R})$, does this mean that the identity operator is not BIBO-stable? Obviously not; this is what we want to clarify here. The argument, which is somewhat technical, rests on the shoulders of two giants: Laurent Schwartz and Lars Hörmander, who were awarded the Fields medal in 1950 and 1962, respectively, for their fundamental contributions to the theory of distributions and partial differential equations.

In the sequel, we shall revisit the topic of BIBO stability with the help of appropriate mathematical tools. In Section II, we recall the classical integral definition of a convolution operator. We then present a correction to the standard characterization of BIBO-stable filters (Proposition 1) together with a new upgraded proof. Since the underlying assumption that the impulse response should be a measurable function excludes the identity operator, we first explain in Section III the extended (distributional) form of convolution supported by Schwartz' kernel theorem (Theorem 1). Based on this formalism, we present two Banach-space extensions of the classical result that should settle the issue: a first one (Theorem 2) that imposes that the result of the convolution be continuous, and a second (Theorem 3) that characterizes the BIBO-stable filters in full generality. The mathematical derivations are presented in Section IV, where we also make the connection with known results in harmonic analysis.

We like to mention a similar clarification effort by Hans Feichtinger, who proposes to limit the framework to convolution operators that are operating on $C_0(\mathbb{R})$ (a well-behaved subclass of bounded functions) in order to avoid pathologies [3]. This is another interesting point of view that is complementary to ours, as discussed in Section IV.

II. BIBO Stability: The Classical Formulation

The convolution of two functions $h, f : \mathbb{R} \to \mathbb{R}$ is the function usually specified by

$$t \mapsto (h * f)(t) \triangleq \int_{\mathbb{R}} h(\tau) f(t-\tau) \mathrm{d}\tau \tag{1}$$

under the implicit assumption that the integral in (1) is well defined for any $t \in \mathbb{R}$. This latter point will be clarified as we develop the mathematics. In particular, this requires that the functions f and h both be measurable[2] in the sense of Lebesgue. Here, instead of designating the continuous-time signal by $f(t)$ and its convolved (or filtered) version by $h(t) * f(t)$, as engineers usually do, we are using the less ambiguous mathematical notations $t \mapsto f(t)$ or $f \in L_p(\mathbb{R})$ and $t \mapsto (h * f)(t)$ or $h * f \in L_\infty(\mathbb{R})$.

If we fix h and consider f as the input signal, then (1) defines a linear shift-invariant (LSI) operator (or system) denoted by $\mathrm{T}_h : f \mapsto h * f$. Its impulse response h is then formally described as $h = \mathrm{T}_h\{\delta\}$, where $\delta \in \mathcal{D}'(\mathbb{R})$ is the Dirac distribution and $\mathcal{D}'(\mathbb{R})$ Schwartz' space of distributions [17]. This interpretation is backed by Schwartz' kernel theorem, as explained in Section III-A.

An important practical requirement for an LSI system is that its response to any bounded input remains bounded. There is one mathematical aspect, however, that makes the formulation of BIBO stability nontrivial in the continuous domain: Depending on the context, the input and output boundedness requirements can be strict, with $\|f\|_{L_\infty} = \|f\|_{\sup} \triangleq \sup_{t\in\mathbb{R}} |f(t)| < \infty$, which arises when the function f is continuous (i.e., $f \in C(\mathbb{R})$), or in the looser sense of Lebesgue: $|f(t)| \leq \|f\|_{L_\infty} < \infty$ for almost any $t \in \mathbb{R}$ (see Section III-B for additional explanations). This latter condition is often expressed as $f \in L_\infty(\mathbb{R})$ where $L_\infty(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \quad \text{s.t.} \quad f \text{ is measurable and } \|f\|_{L_\infty} < \infty\}$ is Lebesgue's space of bounded functions.

The research leading to these results has received funding from the Swiss National Science Foundation under Grant 200020 184646/1.

Biomedical Imaging Group, École polytechnique fédérale de Lausanne (EPFL), Station 17, CH-1015 Lausanne, Switzerland (michael.unser@epfl.ch).

[1] https://en.wikipedia.org/wiki/BIBO_stability. *Accessed November 2019.*

[2] A function $f : \mathbb{R} \to \mathbb{R}$ is said to be Lebesgue-measurable if the preimage $f^{-1}(E)$ of any Borel set E in $\mathbb{R}$ is a Borel set [15]. The property is preserved through pointwise multiplication and translation.

Definition 1. *The linear operator* $\mathrm{T} : f \mapsto \mathrm{T}\{f\}$ *is said to be BIBO-stable if*

1) $\mathrm{T}\{f\}$ *is well-defined for any* $f \in L_\infty(\mathbb{R})$*, and,*
2) *there exists a constant* $C > 0$ *independent of* f *such that*

$$\|\mathrm{T}\{f\}\|_{L_\infty} \leq C\|f\|_{L_\infty}$$

for all $f \in L_\infty(\mathbb{R})$.

The standard condition for the BIBO stability of the continuous-time convolution operator $\mathrm{T}_h : f \mapsto h * f$ that is found in engineering textbooks is $\|h\|_{L_1} < \infty$, where the L_1-norm is defined by

$$\|h\|_{L_1} \triangleq \int_{\mathbb{R}} |h(t)|\mathrm{d}t. \tag{2}$$

A slightly more precise statement is $h \in L_1(\mathbb{R})$, where

$$L_1(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \text{ s.t. } f \text{ is measurable and } \|f\|_{L_1} < \infty\}$$

is Lebesgue's space of absolutely integrable functions.

The sufficiency of the condition $h \in L_1(\mathbb{R})$ is deduced from the standard estimate

$$\left|\int_{\mathbb{R}} h(\tau)f(t-\tau)\,\mathrm{d}\tau\right| \leqslant \int_{\mathbb{R}} |h(\tau)| \cdot |f(t-\tau)|\,\mathrm{d}\tau \leqslant \left(\int_{\mathbb{R}} |h(\tau)|\,\mathrm{d}\tau\right) \|f\|_{L_\infty},$$

which is valid for any $t \in \mathbb{R}$. The convolution integral (1) is therefore well defined if $f \in L_\infty(\mathbb{R})$, which then also yields the classical bound on BIBO stability

$$\|h * f\|_{L_\infty} \leq \|h\|_{L_1}\,\|f\|_{L_\infty} < \infty. \tag{3}$$

By adapting the argument that is used in the discrete-time formulation of BIBO stability, many authors (see Appendix B) affirm that the condition $h \in L_1(\mathbb{R})$ is also necessary. To that end, they apply the convolution system to a "worst-case" signal

$$f_0(t) = \mathrm{sign}\big(h(-t)\big) \tag{4}$$

in order to produce the strongest response at $t = 0$,

$$(h * f_0)(0) = \int_{-\infty}^{+\infty} h(\tau)\mathrm{sign}\big(h(\tau)\big)\,\mathrm{d}\tau = \int_{-\infty}^{+\infty} |h(\tau)|\,\mathrm{d}\tau,$$

which is then claimed to saturate the stability bound (3) with $\|h * f_0\|_{L_\infty} = \|h\|_{L_1}\,\|f_0\|_{L_\infty}$. Unfortunately, this simple reasoning has two shortcomings. First, unlike in the discrete setting, the characterization of what happens at $t = 0$, which is a set of measure zero, is not sufficient to deduce that $\|h * f_0\|_{L_\infty} \geq (h * f_0)(0)$, unless one invokes the continuity of $t \mapsto (h * f_0)(t)$, which is not yet known at this stage (see Theorem 2). Second, one cannot ensure that the Lebesgue convolution integral (1) is well defined for $f_0 \in L_\infty(\mathbb{R})$, unless h is Lebesgue-integrable[3], which then considerably limits the scope of the claim about necessity.

Our first practical fix is an extension of the argumentation to the larger space $L_{1,\mathrm{loc}}(\mathbb{R})$ of measurable functions that are *locally* integrable, meaning that $\int_{\mathbb{K}} |h(t)|\mathrm{d}t < \infty$ over any compact domain $\mathbb{K} \subset \mathbb{R}$. The reassuring outcome, which conforms with the practice in the field, is that one can determine the stability of an LSI system by integrating the absolute value of its impulse response—even if h is not *globally* Lebesgue integrable, as in the case of an increasing and possibly oscillating exponential.

Proposition 1. *If* $h \in L_1(\mathbb{R})$*, then the convolution operator* $f \mapsto h * f$ *defined by* (1) *is BIBO-stable with* $\|h * f\|_{L_\infty} \leq \|h\|_{L_1}\|f\|_{L_\infty}$*. Conversely, if the impulse response* h *is measurable and locally integrable with* $\int_{\mathbb{R}} |h(t)|\mathrm{d}t = \infty$*, then the system is not BIBO-stable, in which case it is said to be unstable.*

Proof. The first statement is a paraphrasing of (3). For the converse part, we assume that $h \in L_{1,\mathrm{loc}}(\mathbb{R})$ with $\int_{\mathbb{R}} |h(t)|\mathrm{d}t = \infty$. Because of the local integrability of h, one can then still rely on the definition of the convolution given by (1), but only if the input function f is bounded and compactly supported. By considering the truncated versions $f_{0,T} = f_0 \cdot \chi_{[-T,T]}$ of the worst-case signal (4), we can therefore determine the maximal value of the output signal as

$$(h * f_{0,T})(0) = \int_{-T}^{+T} h(\tau)\mathrm{sign}\big(h(\tau)\big)\mathrm{d}\tau = \int_{-T}^{+T} |h(\tau)|\mathrm{d}\tau.$$

The additional ingredient is the continuity of $t \mapsto (h * f_{0,T})(t)$ in the neighborhood of $t = 0$ (see Proposition 4 in Appendix D), which allows us to conclude that $(h * f_{0,T})(0) \leq \sup_{t\in\mathbb{R}} |h * f_{0,T}(t)| = \|h * f_{0,T}\|_{L_\infty}$. While the latter quantity is finite for any fixed value of T, we have that $\lim_{T\to\infty}(h * f_{0,T})(0) = \int_{\mathbb{R}} |h(t)|\mathrm{d}t = \infty$, which indicates that the output signal becomes unbounded in the limit. This shows that the underlying system is unstable. □

Another way of obtaining Proposition 1 is as a corollary of Theorem 3 (the complete characterization of BIBO-stable systems) and Proposition 3 in Section IV. The important examples of unstable filters that fall within the scope of Proposition 1 are the systems ruled by differential equations with at least one pole in the right-half complex plane; for instance, $h(t) = \mathbb{1}_+(t)\mathrm{e}^{\alpha t}$ with $\mathrm{Re}(\alpha) \geq 0$ [13]. The derivative operator with $h = \delta'$ and the Hilbert transform with $h(t) = 1/(\pi t)$ are unstable as well (as asserted by Theorem 3), but they fall outside the scope of Proposition 1, the first because δ' is not a function (but a distribution), and the second because the function $1/t$ is not locally integrable—in fact, the impulse response of the Hilbert transform is the distribution "$1/(\pi t)$" that requires the use of a "principal value" for the proper definition of the convolution integral [18].

In the stable scenario, where $h \in L_1(\mathbb{R})$, we are able to characterize the underlying filter by its frequency response

$$\widehat{h}(\omega) \triangleq \mathcal{F}\{h\}(\omega) = \int_{\mathbb{R}} h(t)\mathrm{e}^{-\mathrm{j}\omega t}\mathrm{d}t, \tag{5}$$

which is the "classical" Fourier transform of h. Moreover, the Riemann-Lebesgue lemma ensures that $\widehat{h} \in C_0(\mathbb{R})$ with $\|\widehat{h}\|_{\sup} \leq \|h\|_{L_1}$. We recall that $C_0(\mathbb{R})$ is the Banach space

[3] Any measurable function $h : \mathbb{R} \to \mathbb{R}$ admits a unique decomposition as $h = h^+ - h^-$ with $h^+, h^- : \mathbb{R} \to \mathbb{R}_{\geq 0}$. It is Lebesgue integrable if $\min(\|h^+\|_{L_1}, \|h^-\|_{L_1}) < \infty$ [4].

of continuous and bounded functions that decay at infinity, equipped with the sup-norm.

III. Banach Formulations of BIBO Stability

The classical textbook statements on continuous-time BIBO stability, including our reformulation in Proposition 1, have two limitations. First, they exclude the identity operator with $h = \delta$, as explained in Appendix A. Second, they are often evasive concerning the hypotheses under which the condition $h \in L_1(\mathbb{R})$ is necessary (see Appendix B). In this section, we show how this can be corrected by considering appropriate Banach spaces.

A. Extension of the Notion of Convolution

The scope of our mathematical statements relies on Schwartz' famous kernel theorem [5], [16] which delineates the complete class of linear operators that continuously map $\mathcal{D}(\mathbb{R}) \to \mathcal{D}'(\mathbb{R})$. We recall that $\mathcal{D}(\mathbb{R}) = C_{\rm c}^\infty(\mathbb{R})$ is the space of smooth and compactly supported test functions equipped with the usual Schwartz topology.[4] Its topological dual $\mathcal{D}'(\mathbb{R})$ is the space of generalized functions also known as distributions. In essence, a distribution $f \in \mathcal{D}'(\mathbb{R})$ is a linear map—more precisely, a continuous linear functional—that assigns a real number to each test function $\varphi \in \mathcal{D}(\mathbb{R})$; this is denoted by $f : \varphi \mapsto \langle f, \varphi \rangle$. For instance, the definition of Dirac's impulse as a distribution is $\delta : \varphi \mapsto \langle \delta, \varphi \rangle \triangleq \varphi(0)$.

Beside linearity, the property that defines an LSI operator is $\mathrm{T}_{\rm LSI}\{\varphi(\cdot - t_0)\}(t) = \mathrm{T}_{\rm LSI}\{\varphi\}(t - t_0)$ for any $t_0 \in \mathbb{R}$. Schwartz' theorem then tells us that there is a one-to-one correspondence between continuous LSI operators $\mathcal{D}(\mathbb{R}) \to C(\mathbb{R})$ and distributions, with the defining distribution $h \in \mathcal{D}'(\mathbb{R})$ being the impulse response of the operator. The relevant space of continuous functions here is $C(\mathbb{R})$ with the topology of uniform convergence over compact sets, which involves the system of seminorms $\|f\|_N = \sup_{|t| \le N} |f(t)|, N \in \mathbb{N}$. The latter is an extended functional setup that tolerates arbitrary growth at infinity.

Theorem 1 (Schwartz' kernel theorem for LSI operators)**.** *For any given $h \in \mathcal{D}'(\mathbb{R})$, the operator $\mathrm{T}_h : \varphi \mapsto h * \varphi$ with*

$$t \mapsto (h * \varphi)(t) \triangleq \langle h, \varphi(t - \cdot) \rangle \tag{6}$$

is LSI and continuously maps $\mathcal{D}(\mathbb{R}) \xrightarrow{\text{c.}} C(\mathbb{R})$. Conversely, for every LSI operator $\mathrm{T}_{\rm LSI} : \mathcal{D}(\mathbb{R}) \xrightarrow{\text{c.}} C(\mathbb{R})$, there is a unique $h \in \mathcal{D}'(\mathbb{R})$ such that $\mathrm{T}_{\rm LSI} = \mathrm{T}_h : \varphi \mapsto h * \varphi$ where the convolution is specified by (6).

Then, depending on the decay properties of h, it is generally possible to extend the domain of the convolution operator T_h to some appropriate Banach space according to the procedure described in Section IV. For instance, if $h \in L_1(\mathbb{R})$, then T_h has a continuous extension $L_\infty(\mathbb{R}) \to C_{\rm b}(\mathbb{R}) \subset C(\mathbb{R})$ that coincides with the classical definition given by (1).

[4] A sequence of functions $\varphi_k \in \mathcal{D}(\mathbb{R})$ is said to converge to φ in $\mathcal{D}(\mathbb{R})$ if (i) there exists a compact domain $\mathbb{F}$ that includes the support of φ and of all φ_k; and (ii) $\|\varphi_k - \varphi\|_n \to 0$ for all $n \in \mathbb{N}$, where $\|\varphi\|_n \triangleq \|\mathrm{D}^n \varphi\|_{L_\infty}$ with $\mathrm{D}^n : \mathcal{D}(\mathbb{R}) \to \mathcal{D}(\mathbb{R})$ the nth derivative operator.

Finally, we note that, for the cases where the Dirac impulse δ is in the domain of the extended operator (for instance, when $h \in C(\mathbb{R})$), the distributional definition of the convolution given by (6) yields $\mathrm{T}_h\{\delta\} = h * \delta = h$, which explains the term "impulse response."

B. Banach Spaces of Bounded Functions

In order to investigate the issue of BIBO stability, it is helpful to describe the boundedness and continuity properties of functions via their inclusion in appropriate Banach subspaces of $\mathcal{D}'(\mathbb{R})$. The three relevant function spaces are

$$C_0(\mathbb{R}) \subset C_{\rm b}(\mathbb{R}) \subset L_\infty(\mathbb{R}).$$

The central space consists of the subset of bounded functions that are continuous:

$$C_{\rm b}(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \ \text{ s.t. } f \text{ is continuous and } \|f\|_{\sup} < \infty\}.$$

It is a classical example of Banach space—a complete normed vector space [12]. The smaller space $C_0(\mathbb{R})$, which is also equipped with the sup-norm, imposes the additional constraint that $f(t)$ should vanish at $t = \pm\infty$. It is best described as the completion of $\mathcal{D}(\mathbb{R})$ equipped with the sup-norm, which will have its importance in the sequel. This property is indicated by $C_0(\mathbb{R}) = \overline{(\mathcal{D}(\mathbb{R}), \|\cdot\|_{\sup})}$. The concept is also valid for $L_1(\mathbb{R})$, which can be described as $L_1(\mathbb{R}) = \overline{(\mathcal{D}(\mathbb{R}), \|\cdot\|_{L_1})}$, where the L_1-norm is defined by (2) with $f \in \mathcal{D}(\mathbb{R})$ and the integral being classical—in the sense of Riemann. This completion property applies to $L_p(\mathbb{R}) = \overline{(\mathcal{D}(\mathbb{R}), \|\cdot\|_{L_p})}$ with $p \in [1, \infty)$ as well [4, Proposition 8.17, p. 254], but not for $p = \infty$, which explains the importance of the space $C_0(\mathbb{R})$, which is distinct from $L_\infty(\mathbb{R})$.

In order to properly identify $L_\infty(\mathbb{R})$ as a subspace of $\mathcal{D}'(\mathbb{R})$, we shall exploit the property that the L_∞-norm is the dual of the L_1-norm. We therefore choose to define the L_∞-norm as

$$\|f\|_{L_\infty} \triangleq \sup_{\varphi \in \mathcal{D}(\mathbb{R}):\, \|\varphi\|_{L_1} \le 1} \langle f, \varphi \rangle = \sup_{\varphi \in L_1(\mathbb{R}):\, \|\varphi\|_{L_1} \le 1} \langle f, \varphi \rangle, \tag{7}$$

where the central part of (7) takes advantage of the denseness[5] of $\mathcal{D}(\mathbb{R})$ in $L_1(\mathbb{R})$. This yields a definition that is valid not only for (measurable) functions, but also for all $f \in \mathcal{D}'(\mathbb{R})$. Consequently, we can redefine our target space as

$$L_\infty(\mathbb{R}) = \{f \in \mathcal{D}'(\mathbb{R}) : \|f\|_{L_\infty} < \infty\}, \tag{8}$$

which is readily identified as the topological dual of $L_1(\mathbb{R})$; that is, $L_\infty(\mathbb{R}) = \big(L_1(\mathbb{R})\big)'$ due to the dual specification of the L_∞-norm given by the right-hand side of (7).

While (8) defines $L_\infty(\mathbb{R})$ as a subspace of $\mathcal{D}'(\mathbb{R})$, we can also identify its elements as (bounded) measurable functions $f : \mathbb{R} \to \mathbb{R}$ via the classical association

$$\varphi \mapsto \langle f, \varphi \rangle \triangleq \int_{\mathbb{R}} \varphi(t) f(t) \mathrm{d}t, \tag{9}$$

[5] This means that, for any $f \in L_1(\mathbb{R})$ and any $\epsilon > 0$, there exists a function $\varphi_\epsilon \in \mathcal{D}(\mathbb{R})$ such that $\|f - \varphi_\epsilon\|_{L_1} < \epsilon$. It is a direct consequence of $L_1(\mathbb{R}) = \overline{(\mathcal{D}(\mathbb{R}), \|\cdot\|_{L_1})}$.

where the right-hand side of (9) is a standard Lebesgue integral. Now, the main difference between the sup-norm and the L_∞-norm is that, for $f \in L_\infty(\mathbb{R})$ (identified as a function), the inequality $|f(t)| \leq \|f\|_{L_\infty}$ holds for almost every $t \in \mathbb{R}$. This means that it holds over the whole real line except, possibly, on a set of measure zero. This is often indicated as $\|f\|_\infty = \operatorname{ess\,sup}_{t\in\mathbb{R}} |f(t)|$, using the notion of essential supremum. In other words, the L_∞-norm is more permissive than the sup-norm with $\|f\|_\infty \leq \|f\|_{\sup}$. However, the two norms are equal whenever the function f is continuous, which translates into the isometric inclusion $C_0(\mathbb{R}) \overset{\text{iso.}}{\hookrightarrow} C_\mathrm{b}(\mathbb{R}) \overset{\text{iso.}}{\hookrightarrow} L_\infty(\mathbb{R})$.

C. Extended Results on BIBO Stability

Remarkably, the combination of the two latter function spaces enables us to formulate a first Banach extension of the classical statement on BIBO stability. To that end, we shall restrict the distributional framework covered by Theorem 1 to the case where the impulse response h is identifiable as a measurable function $\big($i.e., $h \in L_{1,\mathrm{loc}}(\mathbb{R}) \subset \mathcal{D}'(\mathbb{R})\big)$. The linear functional on the right-hand side of (6) then has an explicit integral description given by (1) with $f = \varphi \in \mathcal{D}(\mathbb{R})$. Within this class of convolution operators, we now identify the ones whose domain can be extended to $L_\infty(\mathbb{R})$.

Theorem 2. *The convolution operator $\mathrm{T}_h : f \mapsto h * f$ with $h \in L_{1,\mathrm{loc}}(\mathbb{R})$ has a continuous extension $L_\infty(\mathbb{R}) \overset{\text{c.}}{\longrightarrow} C_\mathrm{b}(\mathbb{R})$ if and only if $h \in L_1(\mathbb{R})$. Moreover,*

$$\|h * f\|_{\sup} \leq \|h\|_{L_1} \|f\|_{L_\infty}$$

with the bound being sharp in the sense that it also yields the norm of the underlying operator: $\|\mathrm{T}_h\|_{L_\infty \to C_\mathrm{b}} = \|h\|_{L_1}$ (see Definition 2).

The proof of this result is deferred to Section IV (see Item 2) and the final statement of Theorem 4.

It is of interest to compare Proposition 1 and Theorem 2 because they address the problem of stability from different but complementary perspectives. Proposition 1 is focused primarily on the well-posedness of the convolution integral (1) for $f \in L_\infty(\mathbb{R})$. It can be paraphrased as: "Let h be a measurable (and locally integrable) function. Then, the Lebesgue integral (1) defines a convolution operator that is BIBO-stable if and only if $h \in L_1(\mathbb{R})$." By contrast, Theorem 2 considers the complete family of "classical" convolution operators $\mathrm{T}_h : \mathcal{D}(\mathbb{R}) \to C(\mathbb{R})$ with $h \in L_{1,\mathrm{loc}}(\mathbb{R})$ and precisely identifies the subset of operators that have a continuous extension from $L_\infty(\mathbb{R}) \to C_\mathrm{b}(\mathbb{R})$. Since $C_\mathrm{b}(\mathbb{R})$ is isometrically embedded in $L_\infty(\mathbb{R})$, this is more informative than Proposition 1 because it also tells us that $(h*f)(t)$ is a continuous function of $t \in \mathbb{R}$. In that respect, we note that the requirement that the convolution of any bounded function f be continuous excludes the use of the identity operator with $h = \delta$ at the onset, even if we extend the framework to $h \in \mathcal{D}'(\mathbb{R})$.

To obtain a more permissive characterization of BIBO stability, we need to extend the range of the operator from $C_\mathrm{b}(\mathbb{R})$ to $L_\infty(\mathbb{R})$, which should then also translate into a corresponding enlargement of the class of admissible impulse responses. We shall delineate the latter in a way that parallels our definition of $L_\infty(\mathbb{R})$, with the roles of the L_1- and sup- (or L_∞-) norms being interchanged. To that end, we first define the $\mathcal{M}$-norm as

$$\|f\|_{\mathcal{M}} \triangleq \sup_{\varphi\in\mathcal{D}(\mathbb{R}):\, \|\varphi\|_{\sup}\leq 1} \langle f, \varphi\rangle. \tag{10}$$

This then yields the Banach space

$$\mathcal{M}(\mathbb{R}) = \{f \in \mathcal{D}'(\mathbb{R}) : \|f\|_{\mathcal{M}} < \infty\}, \tag{11}$$

which also happens to be the space of bounded Radon measures[6] on $C_0(\mathbb{R})$. In other words, $\mathcal{M}(\mathbb{R})$ is the topological dual of $C_0(\mathbb{R})$. Moreover, we can invoke the Riesz-Markov theorem to identify $\mathcal{M}(\mathbb{R}) = \big(C_0(\mathbb{R})\big)'$ with the space of bounded signed Borel measures on $\mathbb{R}$ [15]. Concretely, this means that any $h \in \mathcal{M}(\mathbb{R})$ is associated with a unique Borel measure μ_h, which then gives a concrete definition of the linear functional

$$f \mapsto \langle h, f\rangle \triangleq \int_{\mathbb{R}} f(\tau)\mathrm{d}\mu_h(\tau) \tag{12}$$

for any measurable function f, while the total-variation norm of the measure μ_h is given by $\|\mu_h\|_{\mathrm{TV}} \triangleq \int_{\mathbb{R}} \mathrm{d}|\mu_h| = \|h\|_{\mathcal{M}}$ (see Section IV-C). The main point for us is that $\mathcal{M}(\mathbb{R})$ is a superset of $L_1(\mathbb{R})$, with $\|f\|_{\mathcal{M}} = \|f\|_{L_1}$ for all $f \in L_1(\mathbb{R})$. Moreover, we have that $\delta(\cdot - t_0) \in \mathcal{M}(\mathbb{R})$ for any $t_0 \in \mathbb{R}$ with $\|\delta(\cdot - t_0)\|_{\mathcal{M}} = 1$, as can be readily inferred from (10) by considering a non-negative test function that achieves its maximum $\varphi(t_0) = 1$ at $t = t_0$.

For the cases where the impulse response $h \in \mathcal{M}(\mathbb{R})$ is not an L_1 function, we extend our definition of the original (Lebesgue) convolution integral as

$$t \mapsto (h * f)(t) = \langle h, f(t - \cdot)\rangle \triangleq \int_{\mathbb{R}} f(t - \tau)\mathrm{d}\mu_h(\tau), \tag{13}$$

which is the same as (1) when we can write $\mathrm{d}\mu_h(\tau) = h(\tau)\mathrm{d}\tau$, which happens when the corresponding measure μ_h is absolutely continuous[7] with respect to the Lebesgue measure. A standard manipulation then yields that

$$\begin{aligned} |(h * f)(t)| &\leq \int_{\mathbb{R}} |f(t - \tau)|\, \mathrm{d}|\mu_h|(\tau) \\ &\leq \|f\|_{L_\infty} \int_{\mathbb{R}} \mathrm{d}|\mu_h| = \|f\|_{L_\infty} \|h\|_{\mathcal{M}}, \end{aligned} \tag{14}$$

which is the basis for the direct (easy) part of Theorem 3, where the complete class of BIBO-stable systems is identified, including the identity operator.

Theorem 3. *The convolution operator $\mathrm{T}_h : f \mapsto h * f$ with $h \in \mathcal{D}'(\mathbb{R})$ has a continuous extension $L_\infty(\mathbb{R}) \overset{\text{c.}}{\longrightarrow} L_\infty(\mathbb{R})$ if and only if $h \in \mathcal{M}(\mathbb{R})$. Moreover,*

$$\|h * f\|_\infty \leq \|h\|_{\mathcal{M}}\, \|f\|_{L_\infty}$$

with the bound being sharp in the sense that $\|\mathrm{T}_h\|_{L_\infty \to L_\infty} = \|h\|_{\mathcal{M}}$.

[6] We adhere with Bourbaki's nomenclature to distinguish the two complementary interpretations of a measure: either as a continuous linear functional on $\mathcal{D}(\mathbb{R})$ (Radon measure), or as a set-theoretic additive rule that associates a real number to any Borel set of $\mathbb{R}$ (signed Borel measure) [2], [1].

[7] Another way to put it is that h is the Radon-Nikodym derivative of μ_h.

This result, which is also valid in dimensions higher than 1, is known in harmonic analysis [6, p. 140 Corollary 2.5.9], [18] but much less so in engineering circles. It can be traced back to an early paper by Hörmander that provides a comprehensive treatment of convolution operators on L_p spaces [7]. The reminder of the paper is devoted to the proof of the two theorems on BIBO stability and of some interesting variants (see Theorem 4). To that end, we shall rely on Schwartz' powerful distributional formalism which, as we shall see, allows for a rather soft derivation, once the prerequisites have been laid out.

IV. Mathematical Derivations

A. Extension of Convolution Operators

The most general form of a convolution operator backed by Schwartz' kernel theorem (see Theorem 1) is $\mathrm{T}_h : \mathcal{D}(\mathbb{R}) \to C(\mathbb{R}) \hookrightarrow \mathcal{D}'(\mathbb{R})$ with $h \in \mathcal{D}'(\mathbb{R})$, where $\mathrm{T}_h\{\varphi\}$ is defined by (6) for any $\varphi \in \mathcal{D}(\mathbb{R})$. The two complementary ingredients at play there are: (i) the restriction of the domain to $\mathcal{D}(\mathbb{R})$—the "nicest" class of functions in terms of smoothness and decay—and (ii) the extension of the range to $\mathcal{D}'(\mathbb{R})$, which can accommodate an arbitrary degree of growth (polynomial, or even exponential) at infinity. In other words, the theoretical framework is such that it can deal with the very worst scenarios, including unstable differential systems whose impulse response is exponentially increasing.

Then, depending on the smoothness and decay properties of h, it is usually possible to extend the domain of T_h to some Banach space $\mathcal{X} \supseteq \mathcal{D}(\mathbb{R})$ that is continuously embedded in $\mathcal{D}'(\mathbb{R})$, which is denoted by $\mathcal{X} \hookrightarrow \mathcal{D}'(\mathbb{R})$. For this to be feasible, we require that $\|\cdot\|_{\mathcal{X}}$ be a valid norm on $\mathcal{D}(\mathbb{R})$ and that $\mathcal{D}(\mathbb{R})$ be dense in $\mathcal{X}$, which is equivalent to $\mathcal{X} = \overline{(\mathcal{D}(\mathbb{R}), \|\cdot\|_{\mathcal{X}})}$. In other words, $\mathcal{X}$ is the completion of $\mathcal{D}(\mathbb{R})$ equipped with the $\|\cdot\|_{\mathcal{X}}$-norm.

We start by recalling the definition of the norm of a bounded operator.

Definition 2. *Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be two Banach spaces and T a linear operator $\mathcal{X} \to \mathcal{Y}$. Then, the operator is said to be bounded if*

$$\|\mathrm{T}\|_{\mathcal{X}\to\mathcal{Y}} \triangleq \sup_{f\in\mathcal{X}\setminus\{0\}} \frac{\|\mathrm{T}\{f\}\|_{\mathcal{Y}}}{\|f\|_{\mathcal{X}}} < \infty.$$

A direct consequence of Definition 2 is that a bounded operator $\mathrm{T} : \mathcal{X} \to \mathcal{Y}$ continuously maps $\mathcal{X}$ into $\mathcal{Y}$, as indicated by $\mathrm{T} : \mathcal{X} \xrightarrow{c.} \mathcal{Y}$.

Theorem 2 then describes a functional mechanism that allows us to extend an operator initially defined on $\mathcal{D}(\mathbb{R})$. It is a particularization of a fundamental extension theorem in the theory of Banach spaces [14, Theorem I.7, p. 9].

Proposition 2 (Extension of a linear operator)**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be two Banach subspaces of $\mathcal{D}'(\mathbb{R})$ with the additional property that $\mathcal{D}(\mathbb{R})$ is dense in $\mathcal{X}$. Then, the linear operator $\mathrm{T} : \mathcal{D}(\mathbb{R}) \xrightarrow{c.} \mathcal{D}'(\mathbb{R})$ has a unique continuous extension $\mathcal{X} = \overline{(\mathcal{D}(\mathbb{R}), \|\cdot\|_{\mathcal{X}})} \xrightarrow{c.} \mathcal{Y}$ with $\|\mathrm{T}\|_{\mathcal{X}\to\mathcal{Y}} \le C$ if and only if*

$$(i)\quad \mathrm{T}\{\varphi\} \in \mathcal{Y}, \quad \textit{and} \tag{15}$$

$$(ii)\quad \|\mathrm{T}\{\varphi\}\|_{\mathcal{Y}} \le C\|\varphi\|_{\mathcal{X}} \tag{16}$$

for all $\varphi \in \mathcal{D}(\mathbb{R})$ and some constant $C > 0$.

Since a convolution operator $\mathrm{T}_h : \mathcal{D}(\mathbb{R}) \xrightarrow{c.} \mathcal{D}'(\mathbb{R})$ is uniquely characterized by its impulse response $h \in \mathcal{D}'(\mathbb{R})$, the same holds true for its extension $\mathrm{T}_h : \mathcal{X} \xrightarrow{c.} \mathcal{Y}$, which justifies the use of the same symbol. Rather than defining $\mathrm{T}_h\{f\} = h * f$ through a Lebesgue integral as in (1) or (13), we can therefore rely on (6) and define our extended convolution operator $\mathrm{T}_h : \mathcal{X} \to \mathcal{Y}$ through a limit process. Specifically, we pick a Cauchy sequence (φ_n) in $(\mathcal{D}(\mathbb{R}), \|\cdot\|_{\mathcal{X}})$ such that $\lim_{n\to\infty} \varphi_n = f \in \mathcal{X}$. Then, the sequence of functions $(g_n = h * \varphi_n)$ with

$$t \mapsto (h * \varphi_n)(t) = \langle h, \varphi_n(t - \cdot)\rangle \tag{17}$$

is Cauchy in $\mathcal{Y}$ and converges to a limit $g = \lim_{n\to\infty}(h * \varphi_n) \in \mathcal{Y}$, independently of the choice of the φ_n since the space $\mathcal{Y}$ is complete. We now recapitulate this process in the form of a definition.

Definition 3 (Banach extension of a distributional convolution operator)**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be two Banach subspaces of $\mathcal{D}'(\mathbb{R})$ with the additional property that $\mathcal{D}(\mathbb{R})$ is dense in $\mathcal{X}$. When the two conditions in Theorem 2 hold, the unique continuous extension $\mathrm{T}_h : \mathcal{X} \xrightarrow{c.} \mathcal{Y}$ of the convolution operator specified by (17) with $h \in \mathcal{D}'(\mathbb{R})$ is defined by*

$$\mathrm{T}_h : f \mapsto h * f \triangleq \lim_{n\to\infty} (h * \varphi_n) \in \mathcal{Y}, \tag{18}$$

where (φ_n) is any sequence in $\mathcal{D}(\mathbb{R})$ such that $\lim_{n\to\infty} \|f - \varphi_n\|_{\mathcal{X}} = 0$.

Also important for our purpose is the adjoint operator $\mathrm{T}^* : \mathcal{Y}' \to \mathcal{X}'$, which is the unique linear operator such that

$$\langle g, \mathrm{T}\{f\}\rangle_{\mathcal{Y}'\times\mathcal{Y}} = \langle \mathrm{T}^*\{g\}, f\rangle_{\mathcal{X}'\times\mathcal{X}}$$

for any $g \in \mathcal{Y}'$ and $f \in \mathcal{X}'$, where the spaces $\mathcal{X}'$ and $\mathcal{Y}'$ are the duals of the topological vector spaces $\mathcal{X}$ and $\mathcal{Y}$, respectively. If $\mathrm{T} : \mathcal{X} \xrightarrow{c.} \mathcal{Y}$ is bounded with operator norm $\|\mathrm{T}\|$, then the adjoint $\mathrm{T}^* : \mathcal{Y}' \xrightarrow{c.} \mathcal{X}'$ is bounded with $\|\mathrm{T}^*\| = \|\mathrm{T}\|$. In particular, the adjoint of the convolution operator $\mathrm{T}_h : \mathcal{D}(\mathbb{R}) \xrightarrow{c.} \mathcal{D}'(\mathbb{R})$ is $\mathrm{T}_{h^\vee} : \mathcal{D}(\mathbb{R}) \xrightarrow{c.} \mathcal{D}'(\mathbb{R})$, where $h^\vee$ is the time-reversed impulse response such that $\langle h^\vee, \varphi\rangle = \langle h, \varphi^\vee\rangle$, where $\varphi^\vee(t) \triangleq \varphi(-t)$.

We now briefly show how we make use of these two mechanisms to specify the continuous extension $\mathrm{T}_h : L_\infty(\mathbb{R}) \to L_\infty(\mathbb{R})$ with $h \in \mathcal{M}(\mathbb{R})$ (or, $h \in L_1(\mathbb{R})$) that is implicitly referred to in Theorems 2 and 3. The enabling ingredient there is the continuity bound $\|h^\vee * \varphi\|_{L_1} \le \|h^\vee\|_{\mathcal{M}}\|\varphi\|_{L_1}$ (see proof of Theorem 4, Item 4), which also yields $h^\vee * \varphi \in L_1(\mathbb{R})$ for all $\varphi \in \mathcal{D}(\mathbb{R})$. We then apply Definition 3 to specify the unique extension $\mathrm{T}_{h^\vee} : L_1(\mathbb{R}) \xrightarrow{c.} L_1(\mathbb{R})$. An important point for our argumentation is that this (pre-adjoint) convolution operator also has a concrete implementation as

$$t \mapsto (h^\vee * \varphi)(t) = \langle h^\vee, \varphi(t - \cdot)\rangle = \int_{\mathbb{R}} \varphi(\tau + t)\mathrm{d}\mu_h(\tau), \tag{19}$$

which is supported by the same continuity bound with φ now ranging over $L_1(\mathbb{R})$ instead of the smaller space $\mathcal{D}(\mathbb{R})$. The existence and uniqueness of $\mathrm{T}_{h^\vee} : L_1(\mathbb{R}) \xrightarrow{c.} L_1(\mathbb{R})$ then

guarantees the existence and unicity of the adjoint $\mathrm{T}^*_{h^\vee} : L_\infty(\mathbb{R}) \xrightarrow{\text{c.}} L_\infty(\mathbb{R})$. To show that $\mathrm{T}^*_{h^\vee} = \mathrm{T}_h$, we use the explicit representation of $\mathrm{T}_{h^\vee}$ given by (19) with $h^\vee \in \mathcal{M}(\mathbb{R})$ and invoke Fubini's theorem to justify the interchange of integrals in

$$\begin{aligned}\langle \mathrm{T}_{h^\vee}\{f\}, g\rangle &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(\tau + t) \mathrm{d}\mu_h(\tau) \right) g(t) \mathrm{d}t \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} f(x) g(x-\tau) \mathrm{d}\mu_h(\tau) \mathrm{d}x \\ &= \int_{\mathbb{R}} f(x) \left(\int_{\mathbb{R}} g(x-\tau) \mathrm{d}\mu_h(\tau) \right) \mathrm{d}x \\ &= \langle f, \mathrm{T}_h\{g\} \rangle\end{aligned}$$

for any $f \in L_1(\mathbb{R})$ and $g \in L_\infty(\mathbb{R})$. This proves that the original convolution operator defined by (13) coincides with the adjoint of $\mathrm{T}_{h^\vee} : L_1(\mathbb{R}) \xrightarrow{\text{c.}} L_1(\mathbb{R})$, which is also consistent with the property $h = (h^\vee)^\vee$. Since $\mathcal{D}(\mathbb{R}) \subset L_\infty(\mathbb{R})$, we can therefore uniquely identify $\mathrm{T}_h : L_\infty(\mathbb{R}) \xrightarrow{\text{c.}} L_\infty(\mathbb{R})$ as the extension of $\mathrm{T}_h : \mathcal{D}(\mathbb{R}) \to \mathcal{D}'(\mathbb{R})$ that preserves the adjoint relation $\mathrm{T}^*_{h^\vee} = \mathrm{T}_h$.

B. Proof of Banach Variants of BIBO Stability

The Banach spaces of interest for us are $\mathcal{X} = C_0(\mathbb{R}), L_p(\mathbb{R})$ and $\mathcal{Y} = C_0(\mathbb{R}), C_\mathrm{b}(\mathbb{R}), L_p(\mathbb{R})$ with $p \geq 1$.

Theorem 4. *Depending on the functional properties of its impulse response $h \in \mathcal{D}'(\mathbb{R})$, the convolution operator $\mathrm{T}_h : \mathcal{D}(\mathbb{R}) \xrightarrow{\text{c.}} \mathcal{D}'(\mathbb{R})$ defined by (6) admits the following (unique) continuous extensions*[8]

1) *Let $p, q \in (1, \infty)$ be conjugate exponents with $\frac{1}{p} + \frac{1}{q} = 1$. Then, $h \in L_q(\mathbb{R}) \;\Rightarrow\; \mathrm{T}_h : L_p(\mathbb{R}) \xrightarrow{\text{c.}} C_0(\mathbb{R})$ with $\|\mathrm{T}_h\|_{L_p \to C_0} \leq \|h\|_{L_q}$.*
2) *$h \in L_1(\mathbb{R}) \;\Rightarrow\; \mathrm{T}_h : L_\infty(\mathbb{R}) \xrightarrow{\text{c.}} C_\mathrm{b}(\mathbb{R})$ with $\|\mathrm{T}_h\|_{L_\infty \to C_\mathrm{b}} = \|h\|_{L_1}$.*
3) *$h \in \mathcal{M}(\mathbb{R}) \;\Leftrightarrow\; \mathrm{T}_h : C_0(\mathbb{R}) \xrightarrow{\text{c.}} C_\mathrm{b}(\mathbb{R})$.*
4) *$h \in \mathcal{M}(\mathbb{R}) \;\Leftrightarrow\; \mathrm{T}_h : L_1(\mathbb{R}) \xrightarrow{\text{c.}} L_1(\mathbb{R})$.*
5) *$h \in \mathcal{M}(\mathbb{R}) \;\Leftrightarrow\; \mathrm{T}_h : L_\infty(\mathbb{R}) \xrightarrow{\text{c.}} L_\infty(\mathbb{R})$.*

Moreover, the operator norms for Items 3-5, characterized by an equivalence relation, are $\|\mathrm{T}_h\|_{C_0 \to C_\mathrm{b}} = \|\mathrm{T}_h\|_{L_1 \to L_1} = \|\mathrm{T}_h\|_{L_\infty \to L_\infty} = \|h\|_\mathcal{M}$. Finally, under the hypothesis of local integrability $h \in L_{1,\mathrm{loc}}(\mathbb{R})$, the continuity of $\mathrm{T}_h : L_\infty(\mathbb{R}) \xrightarrow{\text{c.}} C_\mathrm{b}(\mathbb{R})$ implies that $h \in L_1(\mathbb{R})$, which is the converse part of Item 2.

Proof:

Item 1. Under the assumption that $h \in L_q(\mathbb{R})$ with $q \geq 1$, we invoke Hölder's inequality

$$|(h * \varphi)(t)| \leq \int_{\mathbb{R}} |h(\tau)| \cdot |\varphi(t-\tau)| \mathrm{d}\tau \leq \|h\|_{L_q} \|\varphi(\cdot - t)\|_{L_p}$$

for any $\varphi \in \mathcal{D}(\mathbb{R})$, which yields the required upper bound $\|h * \varphi\|_{L_\infty} \leq \|h\|_{L_q} \|\varphi\|_{L_p}$. Likewise, by linearity, we get that

$$\begin{aligned}|(h * \varphi)(t) - (h * \varphi)(t - \Delta t)| &= \big| h * \big(\varphi(t) - \varphi(t - \Delta t)\big) \big| \\ &\leq \|h\|_{L_q} \cdot \|\varphi - \varphi(\cdot - \Delta t)\|_{L_p}.\end{aligned}$$

Due to the constraining topology of $\mathcal{D}(\mathbb{R}^d)$, $\lim_{\Delta t \to 0} \|\varphi - \varphi(\cdot - \Delta t)\|_{L_p} = 0$ for any $p \geq 1$, which proves the continuity of the function $t \mapsto (h * \varphi)(t)$. This leads to the intermediate outcome $h * \varphi \in C_\mathrm{b}(\mathbb{R})$ for all $\varphi \in \mathcal{D}(\mathbb{R})$.

If we now replace h by $\phi \in \mathcal{D}(\mathbb{R})$, we readily deduce that $\mathrm{T}_\phi\{\varphi\} = \phi * \varphi$ is compactly supported; hence, $\mathrm{T}_\phi\{\varphi\} \in C_0(\mathbb{R})$ for all $\varphi \in \mathcal{D}(\mathbb{R})$ with $\|\phi * \varphi\|_{L_\infty} \leq \|\phi\|_{L_q} \|\varphi\|_{L_p}$. We then invoke Theorem 2 with $\mathcal{X} = \overline{(\mathcal{D}(\mathbb{R}), \|\cdot\|_p)}$ to deduce that $\mathrm{T}_\phi : L_p(\mathbb{R}) \xrightarrow{\text{c.}} C_0(\mathbb{R})$ for $p \in [1, \infty)$ and $\mathrm{T}_\phi : C_0(\mathbb{R}) \xrightarrow{\text{c.}} C_0(\mathbb{R})$ for any $\phi \in \mathcal{D}(\mathbb{R})$. Since the convolution is commutative, this implies that $\phi * h = h * \phi \in C_0(\mathbb{R})$ for any $h \in L_q(\mathbb{R})$ with $q \in (1, \infty)$ (resp., $h \in C_0(\mathbb{R})$) and $\phi \in \mathcal{D}(\mathbb{R}) \subset L_p(\mathbb{R})$ which, by completion with respect to the $\|\cdot\|_{L_p}$ norm with $p \in (1, \infty)$, gives $\mathrm{T}_h : L_p(\mathbb{R}) \xrightarrow{\text{c.}} C_0(\mathbb{R})$ with $\|\mathrm{T}_h\|_{L_p \to C_0} \leq \|h\|_{L_q}$ (resp., $\mathrm{T}_h : C_0(\mathbb{R}) \xrightarrow{\text{c.}} C_0(\mathbb{R})$ with $\|\mathrm{T}_h\|_{C_0 \to C_0} = \|h\|_{L_1}$).

Item 3. Since $C_\mathrm{b}(\mathbb{R}) \xhookrightarrow{\text{iso.}} L_\infty(\mathbb{R})$, the relevant duality bound there is (14), which yields $\|h * \varphi\|_{L_\infty} \leq \|h\|_\mathcal{M} \|\varphi\|_{L_\infty}$. This allows us to use the same argument as in Item 1 to show that $\mathrm{T}_h\{\varphi\} \in C_\mathrm{b}(\mathbb{R})$ for all $\varphi \in \mathcal{D}(\mathbb{R})$. Since $C_0(\mathbb{R}) = \overline{(\mathcal{D}(\mathbb{R}^d), \|\cdot\|_{L_\infty})}$, we then apply the proven completion technique to specify the unique operator $\mathrm{T}_h : C_0(\mathbb{R}) \to C_\mathrm{b}(\mathbb{R})$ with $\|\mathrm{T}_h\|_{C_0 \to C_\mathrm{b}} \leq \|h\|_\mathcal{M}$. Conversely, let $\mathrm{T}_h : C_0(\mathbb{R}) \xrightarrow{\text{c.}} C_\mathrm{b}(\mathbb{R})$ with operator norm $\|T_h\|_{C_0 \to C_\mathrm{b}} < \infty$. Then, for any $\varphi \in C_0(\mathbb{R})$,

$$(h * \varphi)(0) = \langle h, \varphi^\vee \rangle \leq \|T_h\|_{C_0 \to C_\mathrm{b}} \|\varphi\|_{L_\infty}$$

with $\varphi^\vee \in C_0(\mathbb{R})$ and $\|\varphi^\vee\|_{L_\infty} = \|\varphi\|_{L_\infty}$. By substituting φ for $\varphi^\vee$ and by recalling that $\mathcal{D}(\mathbb{R})$ is dense in $C_0(\mathbb{R})$, we get that

$$\begin{aligned}\sup_{\varphi \in C_0(\mathbb{R}) \setminus \{0\}} \frac{\langle h, \varphi \rangle}{\|\varphi\|_{L_\infty}} &= \sup_{\varphi \in \mathcal{D}(\mathbb{R}) \setminus \{0\}} \frac{\langle h, \varphi \rangle}{\|\varphi\|_{L_\infty}} \\ &= \|h\|_\mathcal{M} \leq \|\mathrm{T}_h\|_{C_0 \to C_\mathrm{b}}, \qquad (20)\end{aligned}$$

which then also proves that the bound is sharp.

Item 4. The key here is the estimate

$$\begin{aligned}\int_{\mathbb{R}} \big|(h * f)(t)\big| \,\mathrm{d}t &\leq \int_{\mathbb{R}} \int_{\mathbb{R}} |f(t-\tau)| \,\mathrm{d}|\mu_h|(\tau) \,\mathrm{d}t \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x)| \mathrm{d}x \right) \mathrm{d}|\mu_h|(\tau) \\ &\qquad \text{(by Fubini)} \\ &= \left(\int_{\mathbb{R}} |f(x)| \mathrm{d}x \right) \left(\int_{\mathbb{R}} \mathrm{d}|\mu_h| \right) \\ &= \|f\|_{L_1} \|h\|_\mathcal{M},\end{aligned}$$

from which we deduce the boundedness of $\mathrm{T}_h : L_1(\mathbb{R}) \to L_1(\mathbb{R})$ with $\|\mathrm{T}_h\|_{L_1 \to L_1} \leq \|h\|_\mathcal{M}$. (The extension technique is essentially the same as in Item 1 with $p = 1$ and $L_1(\mathbb{R}) = \overline{(\mathcal{D}(\mathbb{R}), \|\cdot\|_{L_1})}$.) The converse implication and the sharpness of the bound will be deduced from Item 5 by duality.

Item 5. Since $L_\infty(\mathbb{R}) = \big(L_1(\mathbb{R})\big)'$, the adjoint of $\mathrm{T}_h : L_1(\mathbb{R}) \xrightarrow{\text{c.}} L_1(\mathbb{R})$ is $\mathrm{T}^*_h = \mathrm{T}_{h^\vee} : L_\infty(\mathbb{R}) \xrightarrow{\text{c.}} L_\infty(\mathbb{R})$. The equivalence $h \in \mathcal{M}(\mathbb{R}) \Leftrightarrow h^\vee \in \mathcal{M}(\mathbb{R})$ implies the continuity of $\mathrm{T}_h : L_\infty(\mathbb{R}) \xrightarrow{\text{c.}} L_\infty(\mathbb{R})$ with $\|\mathrm{T}_h\|_{L_\infty \to L_\infty} \leq \|h\|_\mathcal{M} = \|h^\vee\|_\mathcal{M}$. As for the converse implication, we take advantage of the embedding $C_0(\mathbb{R}) \xhookrightarrow{\text{iso.}} L_\infty(\mathbb{R})$, which allows us to reuse the argument of Item 3.

[8] See Definition 3 and accompanying explanations. The bottom line is that the definition of these operators is compatible with the convolution integral (1) or (13) depending on whether h is a function or a Radon measure.

Item 2 and Its Converse. The first part follows from the beginning of the proof of Item 1, the application of the extension principle for $p = 1$, and the commutativity of the convolution integral, which yields $f * h = h * f \in C_{\rm b}(\mathbb{R})$ with $\|h * f\|_{L_\infty} \leq \|h\|_{L_1}\|f\|_{L_\infty}$ for any $f \in L_\infty(\mathbb{R})$. We show that the bound is sharp by applying the convolution operator to the "worst-case" signal f_0 identified in (4). Conversely, let $\mathrm{T}_h : L_\infty(\mathbb{R}) \xrightarrow{\text{c.}} C_{\rm b}(\mathbb{R})$ with $\|T_h\|_{L_\infty \to C_{\rm b}} < \infty$. Taking advantage of the isometric embedding $C_{\rm b}(\mathbb{R}) \xhookrightarrow{\text{iso.}} L_\infty(\mathbb{R})$, we then invoke the equivalence in Item 5 to deduce that $\|h\|_{\mathcal{M}} < \infty$, which implies that $h \in \mathcal{M}(\mathbb{R})$. The announced equivalence then follows from Proposition 3 in Section IV-C. ■

The result in Item 1 is discussed in most advanced treatises on the Fourier transform (e.g., [4, Proposition 8.8, p 241]). We are including it here in a self-contained form—at the expense of a few more lines in the proof of Item 2—because it nicely characterizes the regularization effect of convolution. The equivalences stated in Item 4 and Item 5 are known in the context of the theory of L_p Fourier multipliers [6, Section 2.5], even though the latter does not seem to have permeated to the engineering literature. The equivalence in Item 4 may also be identified as a special instance of Wendel's theorem in the abstract theory of multipliers on locally compact Abelian groups [11, Theorem 0.1.1, p. 2]. Interestingly, the condition $h \in \mathcal{M}(\mathbb{R})$ is also sufficient for the continuity of $\mathrm{T}_h : L_p(\mathbb{R}) \xrightarrow{\text{c.}} L_p(\mathbb{R})$, a claim that is supported by the Young-type norm inequality

$$\|h * f\|_{L_p} \leq \|h\|_{\mathcal{M}}\, \|f\|_{L_p}, \tag{21}$$

which holds for any $f \in L_p(\mathbb{R})$ with $p \geq 1$. However, (21) is only sharp at the two end points $p = 1, +\infty$, in conformity with the statements in Items 4 and 5. In fact, the only other case where the complete class of convolution operators $\mathrm{T}_h : L_p(\mathbb{R}) \xrightarrow{\text{c.}} L_p(\mathbb{R})$ has been characterized is for $p = 2$, with the necessary and sufficient condition being $\widehat{h} \in L_\infty(\mathbb{R})$ (bounded frequency response) [18, Theorem 3.18, p. 28], which is slightly more permissive than the BIBO requirement. Indeed, $h \in \mathcal{M}(\mathbb{R}) \Rightarrow \widehat{h} \in L_\infty(\mathbb{R})$, whereas the reverse implication does not hold.

We like to single out Item 3 in Theorem 4 as the pivot point that facilitates the derivation of the (nontrivial) reverse implications—namely, the necessity of the condition $h \in \mathcal{M}(\mathbb{R})$. While the listed property is sufficient for our purpose, we can refer to a recent characterization by Feichtinger [3, Theorem 2, p. 499] which, in the present context, translates into the refined statement "$h \in \mathcal{M}(\mathbb{R}) \Leftrightarrow \mathrm{T}_h : C_0(\mathbb{R}) \xrightarrow{\text{c.}} C_0(\mathbb{R})$." The additional element there is the vanishing of $(h * f)(t)$ at infinity, which calls for a more involved proof.

While the statement in Item 2 is a special case of Item 5, as made explicit in Section IV-C, the interesting part of the story is that this restriction induces a smoothing effect on the output, ensuring that the function $t \mapsto (h*f)(t)$ is continuous. There is obviously no such effect for the case $h = \delta \in \mathcal{M}(\mathbb{R})$ (identity) or, by extension, $h_{\rm d} = \sum_{n\in\mathbb{Z}} a[n]\delta(\cdot - n) \in \mathcal{M}(\mathbb{R})$ with $\|h_d\|_{\mathcal{M}} = \|a\|_{\ell_1}$, which corresponds to the continuous-time transposition of a digital filter.

C. Explicit Criterion for BIBO Stability

We now show how to determine $\|h\|_{\mathcal{M}}$ (our extended criterion for BIBO stability) under the assumption that $h \in L_{1,\rm loc}(\mathbb{R})$. Any such impulse response can be identified with a distribution by considering the linear form

$$h : \varphi \mapsto \langle h, \varphi\rangle = \int_{\mathbb{R}} h(t)\varphi(t)\mathrm{d}t, \tag{22}$$

which continuously maps $\mathcal{D}(\mathbb{R}) \to \mathbb{R}$. It turns out that the latter is a special instance of a real-valued Radon measure, which is an extended type of measure whose $\|\cdot\|_{\mathcal{M}}$-norm is not necessarily finite.

Definition 4 ([17]). *A distribution $f \in \mathcal{D}'(\mathbb{R})$ is called a real-valued Radon measure if, for any compact subset $\mathbb{K} \subset \mathbb{R}$, there exists a constant $C_{\mathbb{K}} > 0$ such that*

$$\langle f, \varphi\rangle \leq C_{\mathbb{K}} \sup_{t\in\mathbb{K}} |\varphi(t)| \tag{23}$$

for all $\varphi \in \mathcal{D}(\mathbb{K}) = \big\{\varphi \in \mathcal{D}(\mathbb{R}) : \varphi(t) = 0, \forall t \notin \mathbb{K}\big\}$.

A distribution $f^+ \in \mathcal{D}'(\mathbb{R})$ is said to be positive if $\langle f^+, \varphi\rangle \geq 0$ for all $\varphi \in \mathcal{D}^+(\mathbb{R}) = \big\{\varphi \in \mathcal{D}(\mathbb{R}) : \varphi(t) \geq 0, t \in \mathbb{R}\big\}$.

The connection between the two kinds of distributions in Definition 4 is that a positive distribution is a special instance of a Radon measure, while any real-valued Radon measure f admits a unique decomposition as $f = (f^+ - f^-)$, where both $f^+, f^- \geq 0$ are positive distributions [19, Theorem 21.2, p. 218]. One then also defines the corresponding "total-variation measure" $|f| = f^+ + f^-$, which is positive by construction.

It turns out that the Dirac impulse δ is a positive Radon measure with a universal bounding constant $C_{\mathbb{K}} = 1$. Likewise, the minimal constant in (23) for $f \in L_{1,\rm loc}(\mathbb{R})$ is $\mathbb{C}_{\mathbb{K}} = \int_{\mathbb{K}} |f(t)|\mathrm{d}t$, which is essentially what is expressed in Proposition 3.

Proposition 3 (Total-variation norm for measurable functions)**.** *Let $h \in L_{1,\rm loc}(\mathbb{R})$. Then, $\|h\|_{\mathcal{M}} = \|h\|_{L_1} = \int_{-\infty}^{+\infty} |h(t)|\mathrm{d}t$. Consequently, $h \in \mathcal{M}(\mathbb{R})$ if and only if $\int_{-\infty}^{+\infty} |h(t)|\mathrm{d}t < \infty$.*

Proof: In accordance with Definition 4, we view $h \in L_{1,\rm loc}(\mathbb{R})$ as a real-valued Radon measure with $h^+(t) = \max\big(h(t), 0\big)$ and $h^-(t) = \max\big(0, -h(t)\big)$, while the corresponding total-variation measure is $|h| = h^+ + h^- \in L_{1,\rm loc}(\mathbb{R})$ with $|h| : t \mapsto |h(t)|$, which is consistent with the notation. We then distinguish between two cases.

(i) Bounded Scenario. When $\|h\|_{\mathcal{M}} < \infty$, we can invoke the classical Jordan decomposition of a measure (see [4]),

$$\forall f \in \mathcal{M}(\mathbb{R}) : \ \|f\|_{\mathcal{M}} = \|f^+\|_{\mathcal{M}} + \|f^-\|_{\mathcal{M}} = \||f|\|_{\mathcal{M}} < \infty,$$

which allows us to reduce the problem to the easier determination of $\||h|\|_{\mathcal{M}}$. Accordingly, for any given $T > 0$, we define $h_T = |h| \cdot \mathbb{1}_{[-T,T]} \geq 0$ and observe that

$$\begin{aligned}\|h_T\|_{\mathcal{M}} = \sup_{\varphi\in\mathcal{D}(\mathbb{R}):\, \|\varphi\|_{L_\infty}\leq 1} \langle h_T, \varphi\rangle &= \langle h_T, 1\rangle \\ &= \|h_T\|_{L_1} < \infty,\end{aligned}$$

where the supremum is achieved by considering any test function $\varphi_T \in \mathcal{D}(\mathbb{R})$ such that $\varphi_T(t) = 1$ for all $t \in [-T, T]$. In

the limit, we get that $\lim_{T\to\infty}\|h_T\|_{L_1} = \lim_{T\to\infty}\|h_T\|_{\mathcal{M}} = \||h|\|_{\mathcal{M}} < \infty$, from which we conclude that $\||h|\|_{\mathcal{M}} = \|h\|_{\mathcal{M}} = \|h\|_{L_1}$.

(ii) Unbounded Scenario. The condition $\|h\|_{\mathcal{M}} = \infty$ can be formalized as: for any $n \in \mathbb{N}$, there exists $\varphi_n \in \mathcal{D}(\mathbb{R})$ with $\|\varphi_n\|_{L_\infty} = 1$ such that $\langle h, \varphi_n\rangle > n$. However,

$$\langle h, \varphi_n\rangle \leq \Big|\int_{\mathbb{R}} h(t)\varphi_n(t)\mathrm{d}t\Big| \leq \int_{\mathbb{R}} |h(t)|\mathrm{d}t\ \|\varphi_n\|_{L_\infty} = \|h\|_{L_1}.$$

Therefore, $\|h\|_{L_1} > n$ for all $n \in \mathbb{N}$, leading to $\|h\|_{L_1} = \infty$. ■

Let us now conclude with a few more observations.

Since $L_{1,\mathrm{loc}}(\mathbb{R})$ can be identified as the subspace of measures that are absolutely continuous (see [17, p. 18]), the result in Proposition 3 is consistent with the well-known property in probability theory that $L_1(\mathbb{R})$ coincides with the subset of bounded measures that are absolutely continuous.

Under the minimalistic assumption that $h \in L_{1,\mathrm{loc}}(\mathbb{R})$, the convolution integral (1) is well defined for any $t \in \mathbb{R}$ provided that the input function $f : \mathbb{R} \to \mathbb{R}$ is *bounded* and *compactly supported*. Equation (1) then even yields an output function $t \mapsto (h * f)(t)$ that is continuous, as shown in Appendix D. However, the trouble comes from the fact that the output then inherits the potential lack of decay of h when $h \notin L_1(\mathbb{R})$.

One can also make a connection between the result in Proposition 3 and the standard argument that is presented to justify the necessity of $h \in L_1(\mathbb{R})$. When the latter condition is fulfilled, we have that

$$\begin{aligned}\|h\|_{\mathcal{M}} &= \sup_{\varphi\in\mathcal{D}(\mathbb{R}):\,\|\varphi\|_{L_\infty}\leq 1} \langle h, \varphi\rangle \\ &= \|h\|_{L_1} = \sup_{\phi\in L_\infty(\mathbb{R}):\,\|\phi\|_{L_\infty}\leq 1} \langle h, \phi\rangle = \int_{\mathbb{R}} h(t)\phi_0(t)\mathrm{d}t,\end{aligned} \tag{24}$$

where $\phi_0(t) = \mathrm{sign}\big(h(t)\big)$. While the supremum is achieved exactly over $L_\infty(\mathbb{R})$ by taking $\phi = \phi_0$, it is a bit trickier over $\mathcal{D}(\mathbb{R})$ because of the additional smoothness requirement. Yet, due to the definition of the supremum, for any $\epsilon > 0$ there exists a function $\varphi_\epsilon \in \mathcal{D}(\mathbb{R})$ with $\|\varphi_\epsilon\|_\infty = 1$ such that $\int h(t)\varphi_\epsilon(t)\mathrm{d}t = (1-\epsilon)\|h\|_{\mathcal{M}} \leq \|h\|_{\mathcal{M}} = \|h\|_{L_1}$. By taking ϵ arbitrarily small, we end up with φ_ϵ being a "smoothed" rendition of ϕ_0, so that the spirit of the initial argument is retained.

Appendix

A. Is the Dirac Distribution a Member of $L_1(\mathbb{R})$?

Let us start with the historical observation that the eponymous impulse δ is already present in the (early) works of both Fourier and Heaviside [10]. The former, as one would expect, defined it via an "improper" integral (the inverse Fourier transform of "1"), while the latter identified δ as the "formal" derivative of the unit step (a.k.a. the Heaviside function). However, the mathematics for giving a rigorous sense to these identifications were missing at the time; one had to wait for the development Schwartz' distribution theory in the 1950s [17], which already shows that the mere process of obtaining a rigorous definition of δ was far from trivial.

From the pragmatic point of view of an engineer, the title question is at the heart of the matter to understand the scope of Proposition 1, and the source of some confusion, too. Let us start by listing the elements that could suggest that the answer to the question is positive.

1) It is common practice to make liberal use of what mathematicians consider abusive notations; in particular, equations such as $f(t) = \int_{\mathbb{R}} \delta(\tau)f(t-\tau)\mathrm{d}\tau$, which could suggest that $\delta(\tau)$ can be manipulated as if it were a classical function of τ.
2) Dirac's δ has the unit "integral" $\langle \delta, 1\rangle = 1$, which is indicated formally as $\int_{\mathbb{R}} \delta(\tau)\mathrm{d}\tau = 1$. Moreover, $\delta \geq 0$ in the sense that it is a positive distribution (see Definition 4).
3) The Dirac impulse is often described as the limit of $\varphi_n(t) = \frac{n}{\sqrt{2\pi}}\mathrm{e}^{-(tn)^2/2}$ as $n \to \infty$, with $\varphi_n \in \mathcal{S}(\mathbb{R})$. Since $\|\varphi_n\|_{L_1} = 1$ for any $n > 0$, this could suggest that $\lim_{n\to\infty}\|\varphi_n\|_{L_1} = 1$ as well.

In order to convince the reader that the answer to the title question is actually negative, we now refute these intuitive arguments one by one.

1) The explicit description of the Dirac impulse as a centered Gaussian distribution whose standard deviation $\sigma_n = 1/n$ tends to zero suggests that $\delta = \lim_{n\to\infty}\varphi_n$ must be entirely localized at $t = 0$. The best attempt at describing this limit in Lebesgue's world of measurable functions would be
$$p_0(t) = \begin{cases} +\infty, & t = 0 \\ 0, & \text{otherwise,}\end{cases}$$
which is equal to zero almost everywhere. However, since the width of the impulse is zero, we get that $\int_{\mathbb{R}} p_0(t)\mathrm{d}t = 0$, which is incompatible with the property that $\int_{\mathbb{R}} \delta(t)\mathrm{d}t = 1$. This points to the impossibility of representing δ by a function in $L_1(\mathbb{R})$ or even in $L_{1,\mathrm{loc}}(\mathbb{R})$. Strictly speaking, δ is defined as a continuous linear functional on $\mathcal{D}(\mathbb{R})$—or, by extension, $C_0(\mathbb{R})$—which precludes the application of any nonlinear operation (such as $|\cdot|^p$) to it.
2) The generalized Fourier transform of δ is $\mathcal{F}\{\delta\} = 1$, which is bounded, but not decreasing at infinity. If δ was included in $L_1(\mathbb{R})$, this would contradict the Riemann-Lebesgue Lemma, which is equivalent to $\mathcal{F} : L_1(\mathbb{R}) \xrightarrow{\text{c.}} C_0(\mathbb{R})$ with $\|\mathcal{F}\|_{L_1\to C_0} = 1$. By contrast, the inclusion $\delta \in \mathcal{M}(\mathbb{R})$ is compatible with the (dual) continuity property of the Fourier transform $\mathcal{F}^*, \mathcal{F} : \mathcal{M}(\mathbb{R}) \xrightarrow{\text{c.}} L_\infty(\mathbb{R})$ with $\|\mathcal{F}^*\|_{\mathcal{M}\to L_0} = 1$.
3) While the sequence of rescaled Gaussians (φ_n) converges to $\delta \in \mathcal{S}'(\mathbb{R}) \hookrightarrow \mathcal{D}'(\mathbb{R})$ in the (weak) topology of $\mathcal{S}'(\mathbb{R})$ (Schwartz' space of tempered distributions), the problem is that it fails to be a Cauchy sequence in the (strong) norm topology of $L_1(\mathbb{R})$. Hence, there is no guarantee that $\delta = \lim_{n\to\infty}\varphi_n$ stays in $L_1(\mathbb{R})$.

B. Examples of Inaccurate Statements on BIBO Stability

This list is far from exhaustive and not intended to downplay the important contributions of the listed people who are

internationally recognized leaders in the field. Its sole purpose is to illustrate the omnipresence of the misconception in the engineering literature, including in some of the most popular and authoritative textbooks in the theory of linear systems and signal processing.

As a start, one can read in the English version of Wikipedia that a necessary and sufficient condition for the BIBO stability of a convolution operator is that its impulse response be absolutely integrable, formulated as $\int_{\mathbb{R}} |h(\tau)| \mathrm{d}\tau = \|h\|_{L_1} < \infty$. In view of the discussion around Proposition 1, this is only correct if one restricts the scope of the statement to those impulse responses that are Lebesgue-measurable and locally integrable.

Kailath mentions in [9, p. 175] that the equivalence between BIBO stability and $h \in L_1(\mathbb{R})$ is well known, and attributes the result to James, Nichols, and Phillips [8]. It turns out that the pioneers of the theory on control and linear systems were focusing their attention on analog systems ruled by ordinary differentiable equations whose impulse responses are sums of causal exponentials and, therefore, Lebesgue-measurable. Kailath then presents a proof on p. 176 that is essentially the one we used for Proposition 1, except that he neither considers a limit process nor explicitly says that h must be (locally) integrable.

Oppenheim and Willsky discuss the property in [13, p. 113-114]. To justify the BIBO stability of the pure time-shift operator (including the identity), they then present an argument in support of the inclusion of $\delta(\cdot - t_0)$ in $L_1(\mathbb{R})$ (Example 2.13) which, in view of the discussion in Appendix A, is flawed.

Vetterli *et al.* claim in [20, Theorem 4.8, p. 357] that the operator T_h is BIBO-stable from $L_\infty(\mathbb{R}) \to L_\infty(\mathbb{R})$ if and only if $h \in L_1(\mathbb{R})$, a statement that is incompatible with Theorem 3. This can be corrected by limiting the scope of the equivalence as in the statement of Theorem 2.

D. Convolution in the "Unstable" Scenario

Here, we characterize the output of a potentially "unstable" filter when the input signal is compactly supported. The enabling hypothesis is the local integrability of the impulse response.

Proposition 4. *Let $f \in L_\infty(\mathbb{R})$ be compactly supported and $h \in L_{1,\mathrm{loc}}(\mathbb{R})$. Then, the function $t \mapsto (h * f)(t)$ defined by (1) is bounded on any compact set $\mathbb{K} \subset \mathbb{R}$ and continuous; that is, $h * f \in C(\mathbb{R})$.*

Proof. Because of the local integrability of h, the convolution integral (1) is well-defined for any $t \in \mathbb{R}$ with

$$(h * f)(t) = \int_{\mathbb{R}} h(\tau) f(t-\tau) \mathrm{d}\tau = \int_{\mathbb{M}} f(x) h(t-x) \mathrm{d}x$$
(by change of variable)

and

$$|(h * f)(t)| \leq \|f\|_{L_\infty} \int_{\mathbb{M}} |h(t \pm \tau)| \mathrm{d}\tau < \infty,$$

where $\mathbb{M}$ is the smallest symmetric interval such that $f(t) = f(-t) = 0$ for all $t \notin \mathbb{M}$. For any given open bounded set $\mathbb{K} \subset \mathbb{R}$, we then observe that

$$\forall t \in \mathbb{K}: \quad (h * f)(t) = (h_{\mathbb{K}+\mathbb{M}} * f)(t)$$

where $h_{\mathbb{K}+\mathbb{M}} = h \cdot \mathbb{1}_{\mathbb{K}+\mathbb{M}}$ is the restriction of the original impulse response to the set $\mathbb{K}+\mathbb{M} = \{t+\tau : t \in \mathbb{K}, \tau \in \mathbb{M}\}$. Since $h_{\mathbb{K}+\mathbb{M}} \in L_1(\mathbb{R})$, one has that

$$\sup_{t \in \mathbb{K}} |h * f(t)| \leq \|f\|_{L_\infty} \|h_{\mathbb{K}+\mathbb{M}}\|_{L_1} < \infty. \tag{25}$$

Likewise, for any $t, t_0 \in \mathbb{K}$, we have that

$$|h * f(t) - h * f(t_0)| \leq \|f\|_{L_\infty} \|h_{\mathbb{K}+\mathbb{M}}(t - \cdot) - h_{\mathbb{K}+\mathbb{M}}(t_0 - \cdot)\|_{L_1}. \tag{26}$$

Next, we invoke Lebesgue's dominated-convergence theorem and the property that $C_0(\mathbb{R})$ is dense in $L_1(\mathbb{R})$ to show that $\|h_{\mathbb{K}+\mathbb{M}}(t-\cdot) - h_{\mathbb{K}+\mathbb{M}}(t_0-\cdot)\|_{L_1} \to 0$ as $t \to t_0$. This, together with (26), implies that $\lim_{t \to t_0} |(h * f)(t) - (h * f)(t_0)| = 0$, which expresses the continuity of $t \mapsto h * f(t)$ at $t = t_0$ for any $t_0 \in \mathbb{K}$. □

Acknowledgments

The author is extremely thankful to Julien Fageot, Shayan Aziznejad and Hans Feichtinger for having spotted inconsistencies in earlier versions of the manuscript and for their helpful advice. He is also appreciative of Thomas Kailath and Vivek Goyal's feedback, as well as that of the three anonymous reviewers.

References

[1] Jean-Michel Bony. *Cours d'Analyse: Théorie des Distributions et Analyse de Fourier*. Editions Ecole Polytechnique, 2001.

[2] Nicolas Bourbaki. *Elements of Mathematics: 6. Integration*. Springer, 2004.

[3] Hans G. Feichtinger. A novel mathematical approach to the theory of translation invariant linear systems. In *Recent Applications of Harmonic Analysis to Function Spaces, Differential Equations, and Data Science*, pages 483–516. Springer, 2017.

[4] Gerald B. Folland. *Real Analysis: Modern Techniques and their Applications*. John Wiley & Sons, 2013.

[5] I. M. Gelfand and N. Ya. Vilenkin. *Generalized Functions. Vol. 4. Applications of Harmonic Analysis*. Academic Press, New York, USA, 1964.

[6] Loukas Grafakos. *Classical and Modern Fourier Analysis*. Prentice Hall, 2004.

[7] Lars Hörmander. Estimates for translation invariant operators in L_p spaces. *Acta Mathematica*, 104(1):93–140, December 1960.

[8] Hubert M. James, Nathaniel B. Nichols, and Ralph S. Phillips. *Theory of Servomechanisms*, volume 25. McGraw-Hill New York, 1947.

[9] Thomas Kailath. *Linear Systems*, volume 156. Prentice-Hall Englewood Cliffs, NJ, 1980.

[10] Hikosaburo Komatsu. Fourier's hyperfunctions and Heaviside's pseudodifferential operators. In Takahiro Kawai and Keiko Fujita, editors, *Microlocal Analysis and Complex Fourier Analysis*, pages 200–214. World Scientific, 2002.

[11] Ronald Larsen. *An Introduction to the Theory of Multipliers*. Springer-Verlag, Berlin, 1970.

[12] Robert E. Megginson. *An Introduction to Banach Space Theory*. Springer, 1998.

[13] Alan V. Oppenheim and Alan S. Willsky. *Signal and Systems*. Prentice Hall, Upper Saddle River, NJ, 1996.

[14] Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics. Vol. 1: Functional Analysis*. Academic Press, 1980.

[15] Walter Rudin. *Real and Complex Analysis*. McGraw-Hill, New York, 3rd edition, 1987.

[16] Laurent Schwartz. Théorie des noyaux. In *Proc. International Congress of Mathematics 1950*, volume 1, pages 220–230, Providence, RI, 1952. American Mathematical Society.
[17] Laurent Schwartz. *Théorie des Distributions*. Hermann, Paris, 1966.
[18] Elias M. Stein and Guido Weiss. *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton University Press, Princeton, NJ, 1971.
[19] François Trèves. *Topological Vector Spaces, Distributions and Kernels*. Dover Publications, 2006.
[20] Martin Vetterli, Jelena Kovačević, and Vivek K. Goyal. *Foundations of Signal Processing*. Cambridge University Press, 2014.

Michael Unser (M'89–SM'94–F'99) is professor and director of EPFL's Biomedical Imaging Group, Lausanne, Switzerland. His primary area of investigation is biomedical image processing. He is internationally recognized for his research contributions to sampling theory, wavelets, the use of splines for image processing, stochastic processes, and computational bioimaging. He has published over 350 journal papers on those topics. He is the author with P. Tafti of the book "An introduction to sparse stochastic processes", Cambridge University Press 2014.

From 1985 to 1997, he was with the Biomedical Engineering and Instrumentation Program, National Institutes of Health, Bethesda USA, conducting research on bioimaging.

Dr. Unser has served on the editorial board of most of the primary journals in his field including the IEEE Transactions on Medical Imaging (associate Editor-in-Chief 2003-2005), IEEE Trans. Image Processing, Proc. of IEEE, and SIAM J. of Imaging Sciences. He is the founding chair of the technical committee on Bio Imaging and Signal Processing (BISP) of the IEEE Signal Processing Society.

Prof. Unser is a fellow of the IEEE (1999), an EURASIP fellow (2009), and a member of the Swiss Academy of Engineering Sciences. He is the recipient of several international prizes including five IEEE-SPS Best Paper Awards, two Technical Achievement Awards from the IEEE (2008 SPS and EMBS 2010), the Technical Achievement Award from EURASIP (2018), and a recent Career Achievement Award (IEEE EMBS 2020).