DesignCon 2015

Causality Demystified

Stefaan Sercu, Samtec
stefaan.sercu@samtec.com

Chris Kocuba, Samtec

Jim Nadolny, Samtec

Abstract
The quality of time domain simulation results depends on the quality of the S–parameter models used. Causality is shown to be an important parameter limiting the usefulness of a model. In this paper, we demonstrate that causality problems can be classified as mathematical (or numerical) and physical in origin. Mathematical non-causalities are caused by discretization and truncation of S-parameters models, while physical non-causalities are caused by simulations and/or measurement inaccuracies and noise. It will be shown how mathematical non-causalities can be controlled, but not eliminated. Their impact on simulation results will be discussed. Finally, it will be demonstrated that for high quality S-parameters, physical non-causalities can be eliminated easily, while this is not the case if the quality of the S-parameters is not good.

Author(s) Biography

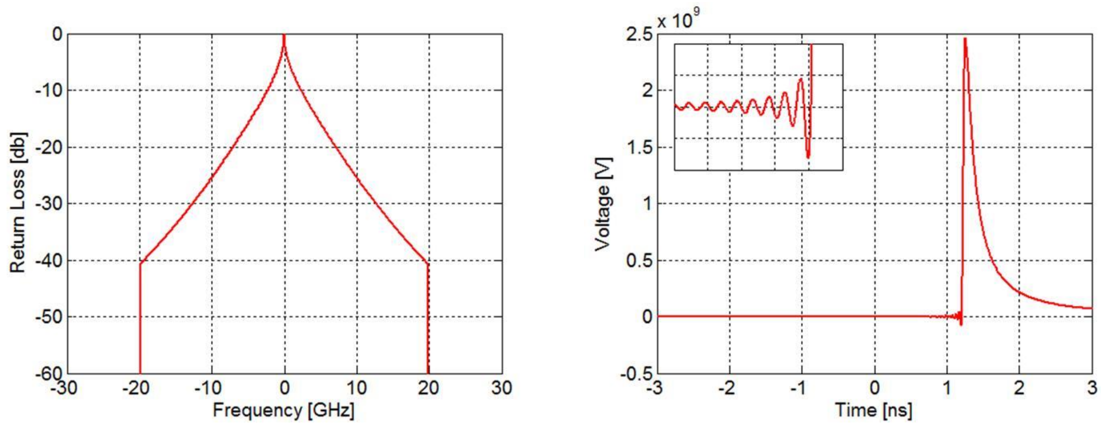
Stefaan Sercu received his MSEE from the University of Ghent, Belgium in 1992. In 1999, he obtained his PhD degree in electrical engineering from the same university. From 1998 until 2013, he was employed by FCI where his work focused on the development and testing of next generation high speed connectors and interconnection technologies. In 2014, Stefaan joined Samtec. In his current position, he is responsible for all SI related issues of high speed connectors and interconnects.

Chris Kocuba received his BA in Mathematics in 2003 and BSEE in 2012 from the Pennsylvania State University. His career began with Samtec as a co-op in 2012 and was hired on directly after graduation. Within the Signal Integrity Group, his focus is primarily on front line customer support which involves information requests, design review, product knowledge and the creation and validation of cable assembly models.

Jim Nadolny received his BSEE from the University of Connecticut in 1984 and an MSEE from the University of New Mexico in 1992. He began his career focused on EMI design and analysis at the system and component levels for military and commercial platforms. For the last 15 years, his focus has shifted to signal integrity analysis of multi-gigabit data transmission systems.

1. Introduction

S-parameters have become the de facto standard for interconnect modeling as they accurately capture impairments such as crosstalk, reflections and loss. For example, resonant behavior in systems is easily seen when working with S-parameters. While there are many advantages to using S-parameters for SI analysis, there are nagging problems associated with using them in time domain simulations. Often it is assumed that the Fourier transform is an analytically precise means of converting from the frequency domain to the time domain. This would be true if the S-parameters were continuous and spanned all frequencies, unfortunately this is not the case. Real world S-parameters are bandwidth limited and sampled so transformation into the time domain will result in non-causal signals. Gibbs Phenomenon is one well known effect which causes a non-causal time domain signal and is due to finite bandwidth of the S-parameter data set. Gibbs Phenomenon is illustrated in Figure 1; the non-causal time domain response is the ringing that occurs prior to the base delay of approximately 1.2 ns.



(a) (b)

Figure 1: Illustration of Gibbs Phenomenon, (a) Typical bandwidth limited insertion loss of a transmission line, (b) Corresponding impulse response with ringing.

Samtec fulfills over 1200 SI support requests per year, and many of these are requests for S-parameter models of connectors. An increasing number of customers have received warnings related to causality violations when using these models in various 3rd party analysis tools. This research was motivated by the need to improve the quality of said connector models, and to guide customers in their use and limitations.

2. Causality

The concept of a causal system is intuitive; the cause must precede the effect. This concept is broad in nature and spans the fields of theology, law, biology and physics. This paper focuses on physics; more specifically on the mathematical aspects of the Kramers-Kronig relations applied to linear time invariant (LTI) systems such as passive interconnects for electronics. The goal of this work is to understand how causality violations arise and what can be done to minimize their impact.

To illustrate, consider the case of a transmission line in which the electrical length is 1 ns. The physical length and velocity of propagation determine the electrical length; for the sake of this simple example it is 1 ns. If a stimulus were applied at the input, the output cannot respond for at least 1 ns. Any output response happening before 1 ns has elapsed, no matter how small, is a causality violation. While this effect may seem minor, bit error rate (BER) calculations of systems are predicated on very accurate calculations of time domain parameters (jitter for example). Causality errors can cascade leading to over/under predictions of system level parameters such as BER or channel operating margin (COM).

Our approach to understanding and ultimately correcting causality violations will be to separate them into numerical and non-physical components. Gibbs Phenomenon is an example of a numerical non-causality. Numerical non-causalities are caused by two separate attributes:

1. Real world S-parameters are bandwidth limited. A typical data set from a vector network analyzer (VNA) might have a bandwidth of 50 MHz to 20 GHz or higher/lower depending on the equipment limitations. The key point here is that it is not infinity; the data sets are bandwidth limited.

2. Real world S-parameters are a sampled data set. Again, a typical data set from a VNA might have a sample every 10 MHz or 1 MHz. The key point is that it is not continuous; it is a discretized data set.

Non-physical components can best be described as “noise” and can occur in measured or modeled data. For example, a full wave simulation of a PCB trace that uses a non-physical dielectric model can result in a causality violation. Another common example occurs when performing a TRL calibration of a VNA measurement. TRL calibration effectively shifts the reference plane of the S-parameter data set and removes loss from test fixtures. If the TRL calibration structures differ from the test fixture, the resultant reference plane shift and/or loss correction may result in a causality violation.

This paper details several aspects of causality. Specifically, Section 3 discusses causality for continuous signals with an infinite bandwidth. It will be shown that for this case, causality is straight forward; can easily be recognized in both the time and frequency domains, and non-causalities can be easily corrected. Section 4 discusses the numerical issues that occur when data is bandwidth limited and discretized. This section also discusses the link between the Fourier Transform and the Discrete Fourier Transform, which itself is used for discrete, bandwidth limited signals. Section 5 deals explicitly with causality for discrete, bandwidth limited signals. It is shown that with the right assumptions, causality for this kind of signal is as straightforward as causality for continuous signals with infinite bandwidth. Definition, detection and enforcement are also discussed. Section 6 covers non-physical; non-causalities and final conclusions are located in Section 7.

3. Causality for continuous functions with infinite bandwidth

While real world S-parameters are discrete and bandwidth limited, one should first consider an ideal case with S-parameters that are continuous and have an infinite bandwidth. It is well known that in this case, real and imaginary parts of the S-parameters are linked through Kramers-Kronig or the Hilbert transform [1-2]. This section illustrates what these relations represent in the time domain [3]. These very relations will be used as the basis to define associations between the real and imaginary part of S-parameters that are bandwidth limited and discretized. This approach should (hopefully) "demystify" the mathematics by starting with something a bit more basic.

Consider a causal time domain signal v(t) shown in Figure 2.

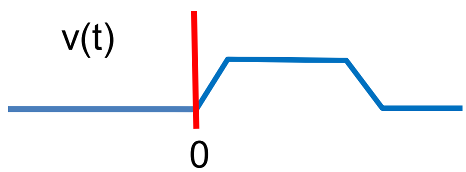


Figure 2: Illustration of a causal time domain signal.

The time domain definition of causality is quite simple:

$$v(t) = 0 \quad t < 0 \tag{1}$$

There can be no signal as long as there is no excitation. This equation assumes that the signal has no delay and that the reference plane is at t=0. This assumption is maintained throughout this document. This is not a limitation; as for passive interconnection structures, it is always possible to decompose the S-parameters into the product of a minimum-phase function and an all pass-function: with the minimum phase function having its reference plane at t=0. An all pass-function is a function with unit magnitude response at all frequencies and is a pure phase shifter [4].

In order to obtain a frequency domain definition of causality, one must first expand v(t) into even and odd components and introduce the sign function. This construct will also be helpful to show how to enforce causality. Recall that an even and odd signal are defined by

$$\begin{aligned} v_{even}(t) &= \frac{v(t) + v(-t)}{2} \\ v_{odd}(t) &= \frac{v(t) - v(-t)}{2} \end{aligned} \tag{2}$$

When taking into account equation (1) (v(t) is causal), it follows:

$$v_{even}(t) = \begin{cases} \dfrac{v(-t)}{2} & t < 0 \\ 0 & t = 0 \\ \dfrac{v(t)}{2} & t > 0 \end{cases} \quad \text{and} \quad v_{odd}(t) = \begin{cases} -\dfrac{v(-t)}{2} & t < 0 \\ 0 & t = 0 \\ \dfrac{v(t)}{2} & t > 0 \end{cases} \tag{3}$$

These equations become

$$\begin{aligned} v_{even}(t) &= \text{sign}(t).v_{odd}(t) \\ v_{odd}(t) &= \text{sign}(t).v_{even}(t) \end{aligned} \tag{4}$$

Through the use of the sign function definition:

$$\text{sign}(t) = \begin{cases} -1 & t < 0 \\ 0 & t = 0 \\ 1 & t > 0 \end{cases} \tag{5}$$

Equation (4) can be considered as a second "definition" of causality. A signal v(t) is causal if the even and odd part of v(t) are linked through the sign function.

Remember, the goal is to provide S-parameters that are of a high quality so as to be useful for SI simulations. S-parameters are frequency domain based; therefore having the causality condition confirmed in the time domain is not very helpful. The next step then becomes to represent the causality conditions in equation (3) using the Fourier Transform. Going forward, "Fourier Transform" will be abbreviated in this paper as "FT". To obtain the causality requirement in the frequency domain, first perform the Fourier Transform of equation (4) and take into account that

$$\begin{aligned} V(f) &= FT(v(t)) = V_R(f) + jV_I(f) \\ V_{even}(f) &= FT(v_{even}(t)) = V_R(f) \\ V_{odd}(f) &= FT(v_{odd}(t)) = jV_I(f) \\ SIGN(f) &= FT(sign(t)) = \frac{1}{j\pi f} \end{aligned} \tag{6}$$

Thus obtaining

$$\begin{aligned} V_R(f) &= SIGN(f) * jV_I(f) = \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{V_I(f')}{f - f'} df' \\ jV_I(f) &= SIGN(f) * V_R(f) = -\frac{j}{\pi} P \int_{-\infty}^{+\infty} \frac{V_r(f')}{f - f'} df' \end{aligned} \tag{7}$$

The integral is defined according to the Cauchy Principle Value:

$$P\int_{-\infty}^{+\infty} = \lim_{\varepsilon \to 0+}\left[\int_{-\infty}^{f-\varepsilon} + \int_{f+\varepsilon}^{+\infty}\right] \tag{8}$$

By taking the Fourier Transform of equation (4), one obtains the Hilbert Transform which relates real and imaginary parts of a causal S-parameter. Translation of the Hilbert Transform for causal functions in the time domain indicates that the even and odd parts of a causal signal are linked with each other.

If an S-parameter is not causal, the above equations also provides a very simple method to enforce causality both in the time and frequency domains. This is illustrated in Figure 3 for the time domain. In the first step, the even (or odd) function is calculated (Figures b, c). Next, the corresponding causal, odd (or even) functions are calculated using equation (4) (Figures d, e). The causality enforced, time domain response is then obtained by taking the sum of the even (or odd) function and the calculated odd (or even) function (Equation (9) and Figure 3f and 3g).

$$\begin{gathered} v_{causal,even}(t) = v_{even}(t) + sign(t).v_{even}(t) \\ v_{causal,odd}(t) = v_{odd}(t) + sign(t).v_{odd}(t) \end{gathered} \tag{9}$$

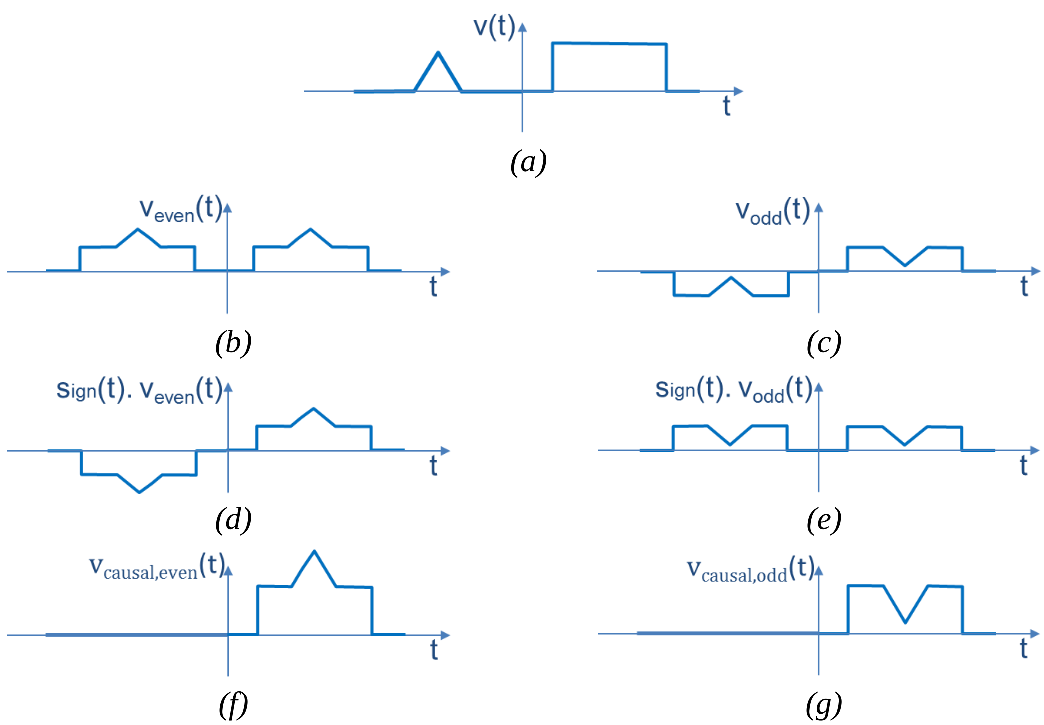


Figure 3: Causality enforcement illustrated.

Notice in Figure 3 that in the causality enforced signals, the non-causality is shifted from negative time to positive time. This is an undesired side-effect of causality enforcement. To avoid this, the even and odd causality enforced functions need to be combined:

$$v_{causal}(t) = \frac{v_{causal,even}(t) + v_{causal,odd}(t)}{2} \tag{10}$$

as illustrated in Figure 4.

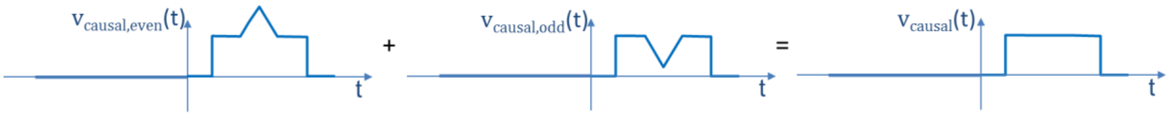


Figure 4: Combination of odd and even causality enforced functions.

By substituting equation (9) in equation (10) we obtain

$$v_{causal}(t) = \frac{v(t)(1+\mathrm{sign}(t))}{2} = u(t).v(t) \tag{11}$$

with u(t) being a perfect step signal. The causality enforced function can simply be obtained by multiplying the original causal function with a perfect step function.

In the frequency domain, these equations become

$$\begin{aligned} V_{causal}(f) &= \frac{V(f) + \mathrm{SIGN}(f) * V(f)}{2} \\ &= \frac{V(f) + \mathrm{Hilbert}(V(f))}{2} \end{aligned} \tag{12}$$

At this point, we introduce the concept of a causality number. The objective is to have a metric which answers the question: "How causal is an S-parameter set?". The metric for signals with an infinite bandwidth is simple and logical. This is done by determining how much energy there is between the original function and the causality enforced function. In the time domain, this is identical to verifying how much energy exists for t<0 as shown is equation (13).

$$\text{Causality Number} = 100 \frac{\int_{-\infty}^{+\infty} \left(v(t) - v_{causal}(t)\right)^2 dt}{\int_{-\infty}^{+\infty} v^2(t)\, dt} = 100 \frac{\int_{-\infty}^{0} v^2(t)\, dt}{\int_{-\infty}^{+\infty} v^2(t)\, dt} \tag{13}$$

The smaller this number, the more causal the resultant signal. This equates to zero for causal functions. In frequency domain, it follows:

$$\text{Causality Number} = 100. \frac{\int_{-\infty}^{+\infty} \left|V(f) - V_{causal}(f)\right|^2 df}{\int_{-\infty}^{+\infty} \left|V(f)\right|^2 df} = 100. \frac{\int_{-\infty}^{+\infty} \left|\Delta V(f)\right|^2 df}{\int_{-\infty}^{+\infty} \left|V(f)\right|^2 df} \tag{14}$$

4. Numerical representation of S-parameters

To process signals, numerical tools cannot work with infinite continuous signals; therefore, the infinite signals must be truncated and discretized. Furthermore, time and frequency domain representations of the signals are linked through the Discrete Fourier Transform (DFT) instead of the Fourier Transform. Unfortunately, non-causality effects are introduced if this is not done with care.

Figure 5 below compares the impulse response of an infinite continuous signal with the impulse response of a bandwidth limited, discretized signal of the same system; they do not fall on top of each other. Before discussing causality for discrete signals, one must understand why there is a difference.

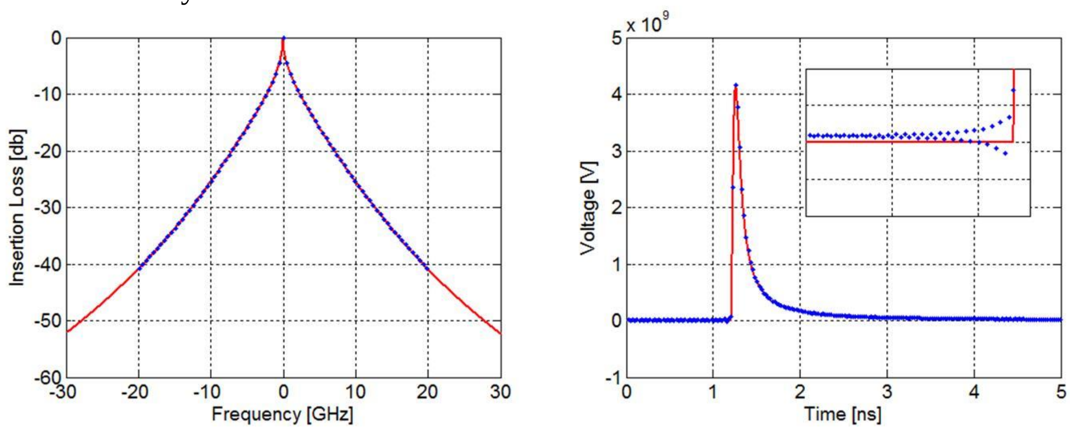


Figure 5: Comparison between the IFT and DIFT of a continuous and sampled, bandwidth limited S-parameter.

What follows is the impact of bandwidth limitation, discretization and the use of the DFT on causality in three steps. First, we detail the effect of bandwidth limitations. Second, we look at the impact of sampling in the frequency domain. Third, we consider sampling in the time domain. As a conclusion, the relation between the DFT and Fourier Transform will be illustrated, followed by a closer look at the Gibbs Phenomenon for sampled bandwidth limited signals.

4.1 Bandwidth limitation of S-parameters

A bandwidth limited signal $H_B(f)$ can be represented as a bandwidth unlimited signal H(f) multiplied with a perfect rectangular filter R(f).

$$H_B(f)=H(f).R(f) \tag{15}$$

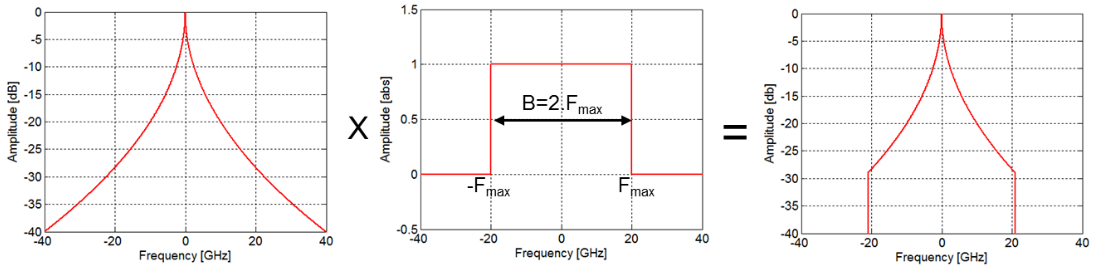


Figure 6: Bandwidth limited signal.

In the time domain, this equation becomes:

$$h_B(t)=B.h(t)*\mathrm{sinc}(\pi Bt) \tag{16}$$

as the inverse Fourier Transform of a block function is the sinc function.

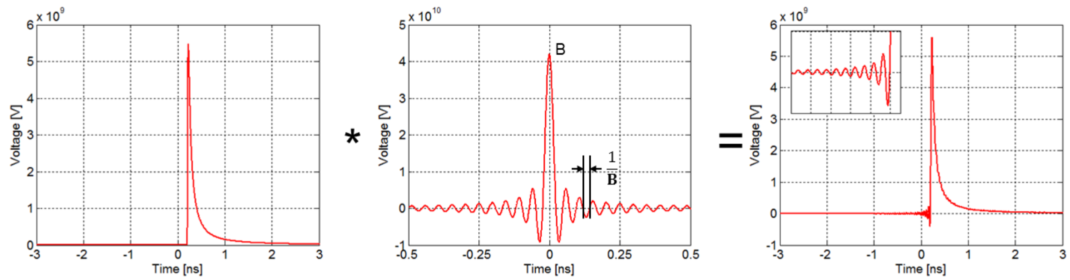
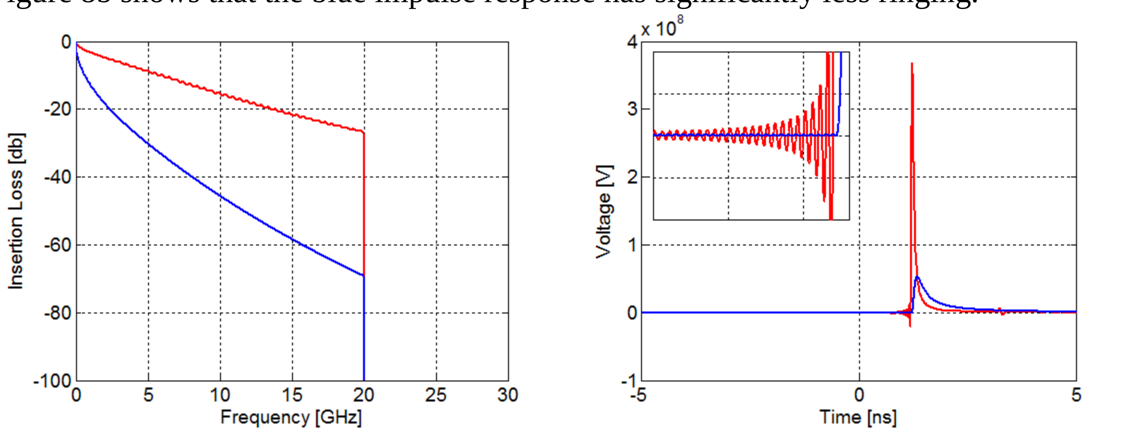


Figure 7: Time domain response bandwidth limited signal

The time domain impulse response of a bandwidth limited signal is equal to the impulse response of the bandwidth unlimited function convolved with a non-causal sinc function. As a result, there is ringing, also known as Gibbs Phenomenon, on the impulse response rendering the impulse response no longer causal. This ringing is a function of the S-Parameter bandwidth and can only be avoided if all of the signal energy falls completely in the considered bandwidth, $F_{max,}$ as is illustrated in Figure 8. Figure 8a compares the spectrum of two different signals. The blue signal has more loss, and as such has relatively more energy (compared to the total energy), in the considered bandwidth. Figure 8b shows that the blue impulse response has significantly less ringing.



(a) *(b)*

Figure 8: Ringing for bandwidth limited S-parameters with different loss: the blue curve has significantly less ripple than the red curve.

Unfortunately, this requirement is quite impractical for many systems (e.g. low loss systems) as this means the maximum frequency must be several hundred gigahertz or higher.

One could consider causality enforcement to minimize ringing by using equations (10) or (12). An example is illustrated in Figure 9. Enforcing causality results in S-parameters that are no longer bandwidth limited. Fortunately, as is shown in paragraph 4.4, there are a number of ways to reduce the ringing. Also, in practical simulations one is not interested in the impulse response, but in the pulse response. To obtain the pulse response, the S-parameters need to be multiplied with the spectrum of the excitation signal: a pulse in this case. Since a pulse can be considered a low pass filter, all energy of the received signal shifts to lower frequencies, and the received signal has relatively more energy in the considered bandwidth, thus reducing the ringing of the pulse response significantly.

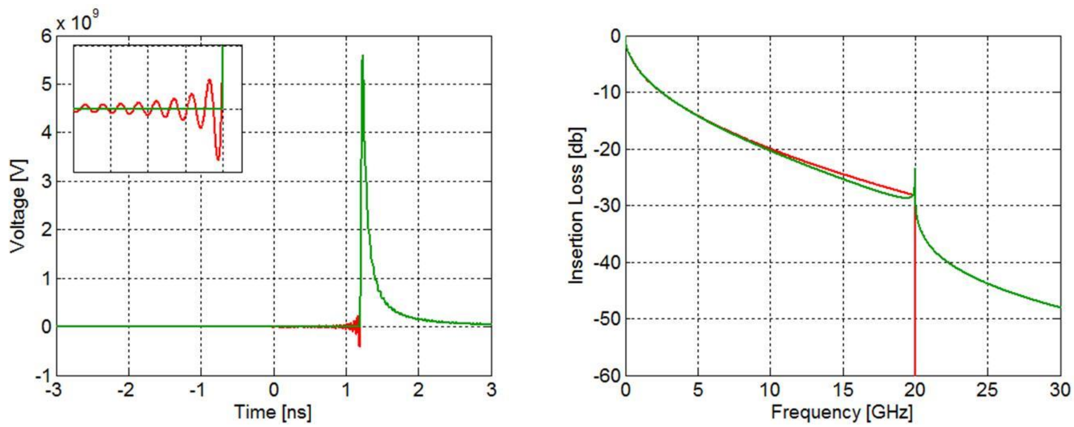


Figure 9: Elimination of ringing by causality enforcement.

4.2 Sampling of S-parameters

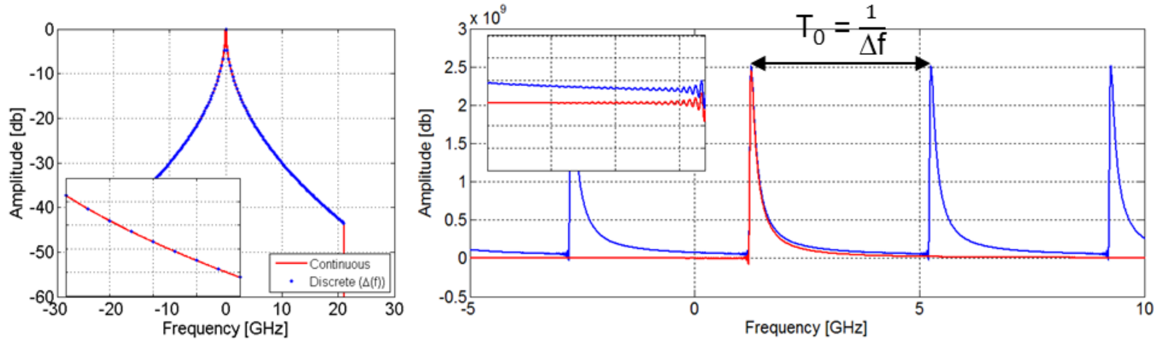
The sampling of S-parameters results in a periodic impulse response with period T_0 defined by the sampling frequency: $T_0 = 1/\Delta f$. This is illustrated in Figure 10. It is assumed that:

- the frequency is continuous and infinite
- the S-parameter data is zero at all frequencies except for the sampling frequencies

The inverse Fourier Transform is used to obtain the time domain data. The impulse response of the sampled bandwidth limited S-parameters is given by

$$h_{discrete}(t) = \sum_{n=-\infty}^{+\infty} h_{continuous}(t - nT_0) \quad n = \ldots,-1,0,1,\ldots \tag{17}$$

With $h_{continuous}(t)$ being the impulse response of the un-sampled bandwidth limited S-parameters.



(a) *(b)*

Figure 10: Sampling of the S-parameters (a) results in a periodic impulse response (b).

Notice that the impulse response of the sampled S-parameters does not equal the sampled impulse response of the continuous S-parameters; there is a shift/offset (see small window in Figure 10b This is caused by time domain overlap or time domain leakage which occurs when the duration of the un-sampled impulse response lasts longer than the period T_0. To avoid this time domain overlap, one must reduce the frequency step (Δf) to make sure that T_0 is larger than the duration of the impulse response as shown in Figure 11.

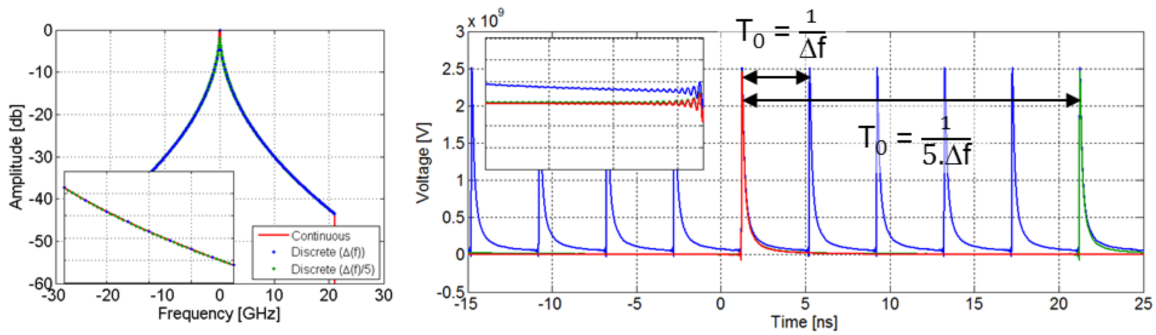
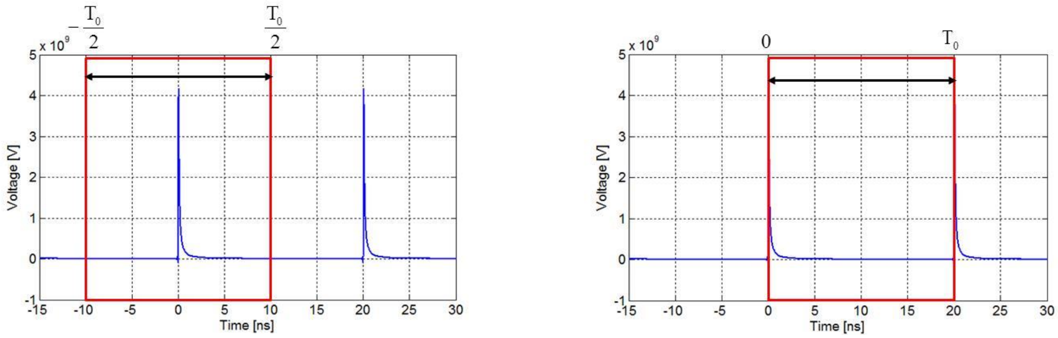


Figure 11: By reducing the frequency step Δf the time window gets larger and the difference between the continuous impulse response and discrete impulse response disappears.

Due to the periodicity, the impulse response is definitely not-causal ($h_{discrete}(t) \neq 0$ for $t<0$). The initial definition for causality can no longer be used.

Since the time domain response is periodic, it can be limited to one period. The question arises, where must the time domain window be taken? For simplicity, the assumption is made that the reference plane is at t=0. It is also assumed that the impulse response falls completely in the time window. Basically, there are 2 options (see Figure 12): a time window from $-T_0/2$ to $T_0/2$ (option a) or a time window from 0 to T_0 (option b).



(a) *(b)*

Figure 12: Options for time window selection.

The answer depends on the assumptions made. If it is known that the S-parameters are causal and do not suffer from physical non-causalities, then option (b) is the best choice: the impulse response can last a complete time window before time domain leakage occurs; this is twice as long as for option a. If it is unknown that the S-parameters are causal, then option (a) must be selected. This is the only option with "negative" time and allows one to check if a signal is causal. Consequently, if the impulse response is longer than half the period, T_0, then according to option (a), the S-parameters become non-causal even if the impulse response is smaller than T_0! High quality S-parameters require that option (a) is selected and that the impulse response falls completely in half the time window.

In order to avoid non-causalities that are a result of time domain overlap or leakage, the time domain period, T_0, must be larger than two times the impulse response duration, T_{max}. The impulse response duration of a typical connector, cable or backplane is determined by several factors: length of delay, impedance mismatches and reflection and insertion loss. Unfortunately, many of the signals we consider have an infinite impulse response duration. Figure 13 shows an RC-filter; the infinite impulse response of this filter is given by:

$$s_{2,1}(t) = \delta(t) - \frac{1}{2Z_0C} e^{-\frac{R+2Z_0}{2Z_0RC}t} u(t) \qquad (18)$$

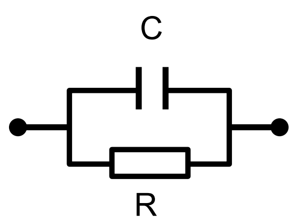


Figure 13: RC filter

The impulse response of a signal propagating through an ideal transmission line with reflection coefficient ρ and delay τ is given by (Figure 14)

$$h(t) = \left(1-\rho^2\right)\left(\delta(t-\tau) + \rho^2\,\delta(t-3\tau) + \rho^4\,\delta(t-3\tau) + \ldots\right) \qquad (19)$$

The impulse response gets smaller as time passes, but is only zero for $t=\infty$.

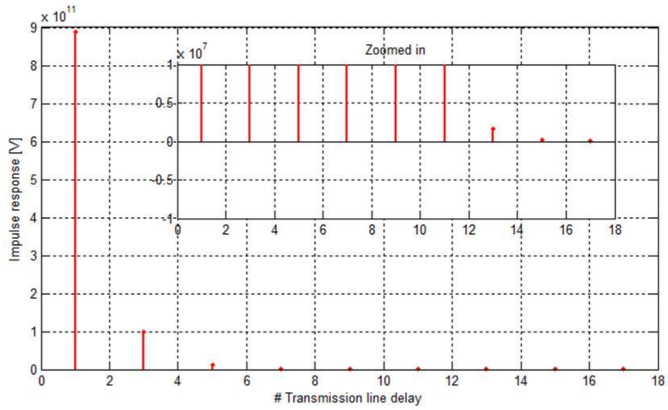
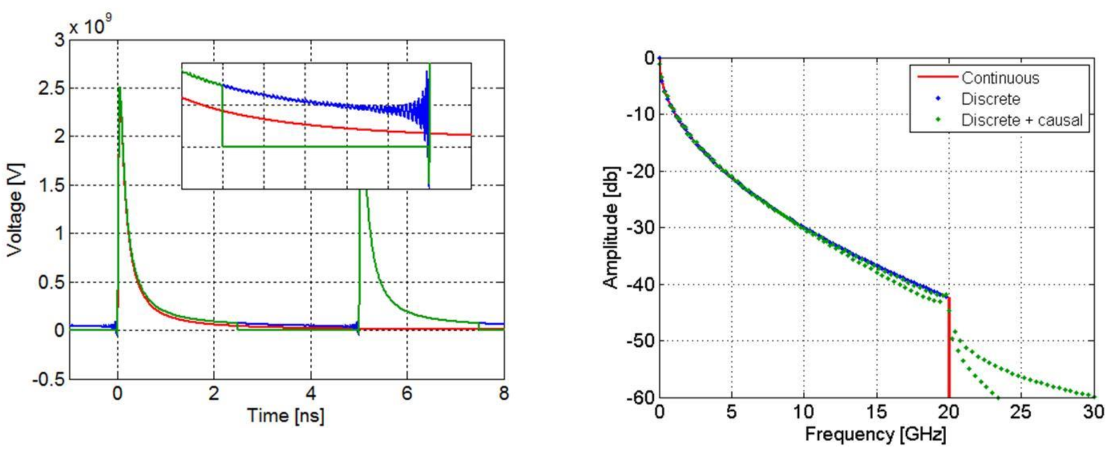


Figure 14: Impulse response of ideal transmission line.

By definition, many signals become non-causal due to discretization. One solution is to minimize and control the non-causalities by making the time window, $T_{0,}$ as large as needed by minimizing the frequency step, Δf. A second method is to limit the impulse duration by truncating the time domain response so it falls completely in half of the time window as is illustrated below. However, applying this method results in S-parameters that are no longer bandwidth limited.

In order to enforce causality, the impulse response is set to zero for the negative time portion of the time window for all periods. The result is a small change to the insertion loss which is shown in Figure 15. The observed change is a function of the non-causality; the more causal the impulse response, the less of a change will be seen upon causality enforcement. If there is no change, then the signal is already causal.



(a) *(b)*

Figure 15: Causality enforcement (a) time domain, (b) frequency domain.

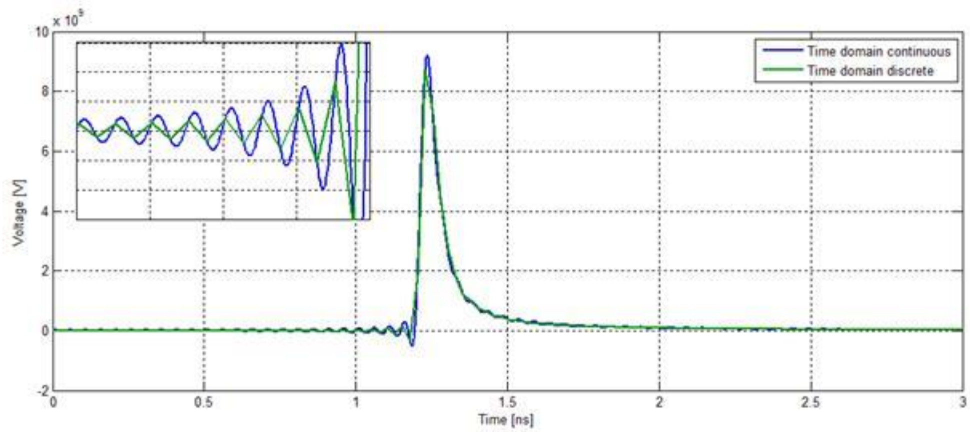
4.3 Sampling time domain data

In the previous sections, we assumed that the bandwidth of frequency domain data is infinite and that time domain signals are continuous. To come to a numerical representation of the impulse response, the time domain response needs to be discretized. In principle, the time step, (Δt), can be chosen arbitrarily, but we will consider a specific case which illustrates how the Fourier Transform correlates with the DFT. We choose the time step according to Equation (20) shown below.

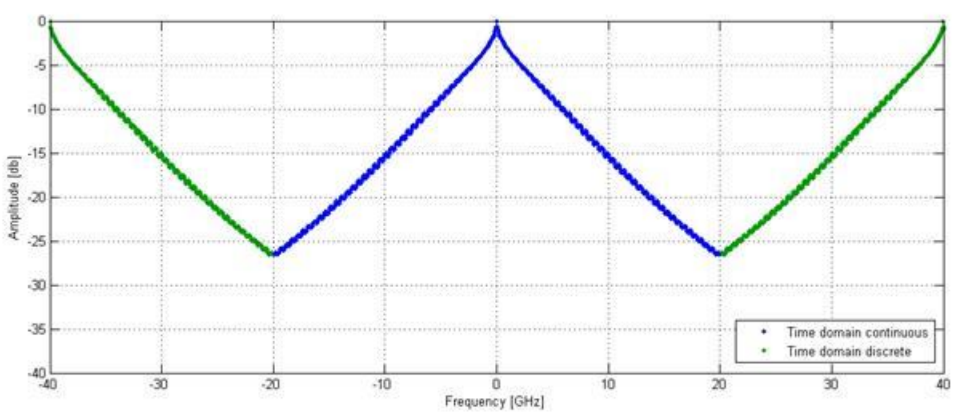
$$\Delta t = \frac{1}{2 \cdot F_{\max}} \tag{20}$$

Figure 16(a) shows the sampled time domain response of an impulse response using a time step, Δt=25ps. Figure 16 (b) shows the corresponding spectrum. As a result of the time domain sampling, the spectrum becomes periodic and has infinite bandwidth.

The frequency domain representation of the finite length, sampled impulse response is periodic and can be limited to one period. Notice that a bandwidth limited set of S-parameters is obtained if we have a continuous sampling of the time domain impulse response. The sampled time domain function, limited to one time window, is the function that is obtained when an Inverse Discrete Fourier Transform (IDFT) is performed on the sampled, bandwidth limited frequency domain data. If the correct assumptions are made,



(a)



(b)

Figure 16: Sampled time domain data and corresponded spectrum.

the IFT and IDFT will result in the same time domain signal. For the IFT, all non-sampled, frequency domain data needs to be set to zero and periodically extended. Notice that the time domain sampling has no impact on causality.

The IDFT is identical to the IFT assuming:

- the spectrum is zero between sampling points
- the signal is periodic with period, T = 2·Fmax
- the time domain response is limited to one period T

The DFT is identical to the FT assuming:

- the signal is zero between sampling points
- the signal is periodic with period, T = 1/Δf
- the frequency spectrum is limited to one period

This indicates that bandwidth limitations and discretization introduce non-causalities that can be limited by maximizing the bandwidth of the signal and minimizing the frequency step.

4.4 Gibbs Phenomenon for discrete bandwidth limited signals

The previous section (Figure 16) showed that Gibbs Phenomenon occurs for bandwidth limited signals, and that the bandwidth limited signal is, when considering the DFT, not bandwidth limited but periodic. This section takes a closer look at the periodic function. Figure 17 shows the real and imaginary parts of the periodic S-parameters. It is clear that the imaginary part is discontinuous at F_{max}= 20 GHz (which will likely be the case for most signals). This discontinuity can be avoided if the S-parameters are limited to 19.8 GHz (Figure 18). This is the first frequency below F_{max} where the imaginary part of the S-parameter is equal to zero. Figure 19 compares the impulse responses of both sets of S-parameters. The ringing is significantly reduced if the S-parameter set has no discontinuities at the maximum frequency.

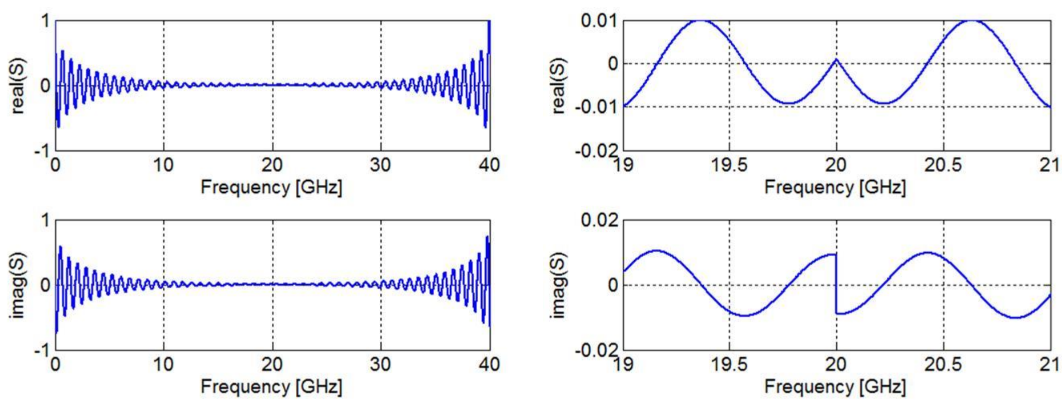


Figure 17: Real and imaginary parts of periodic transfer function

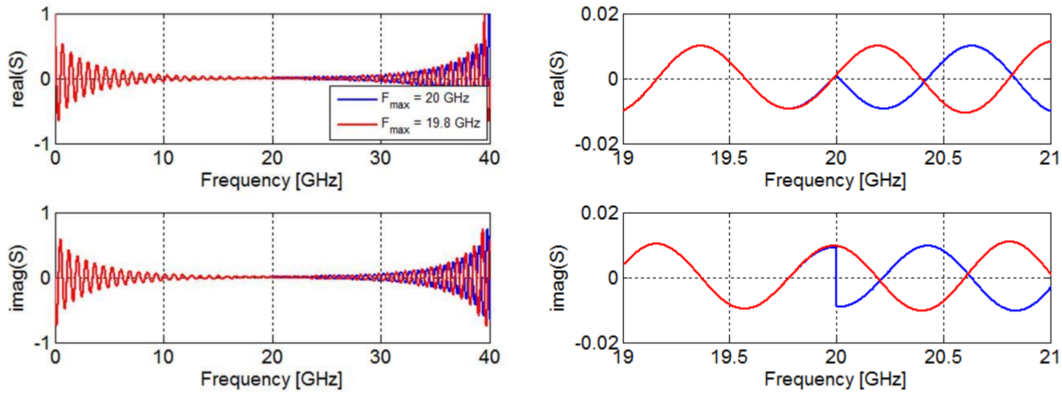


Figure 18: Transfer function made periodic by reducing the bandwidth from 20 GHz to 19.8 GHz.

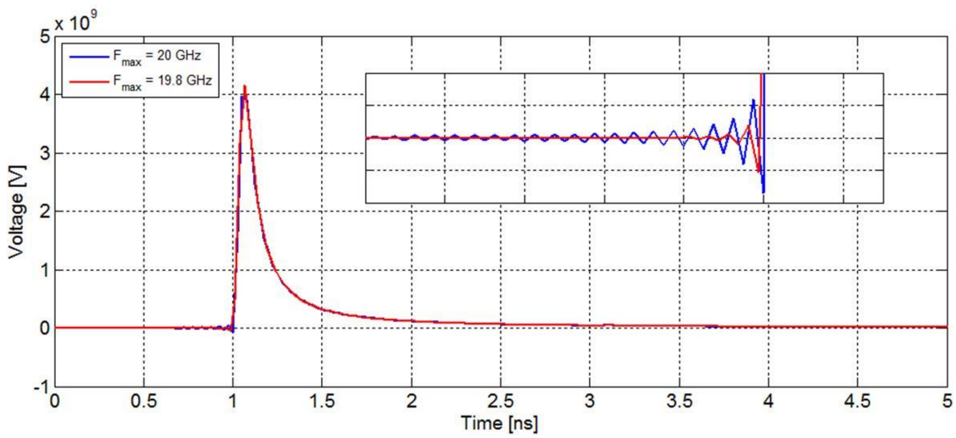
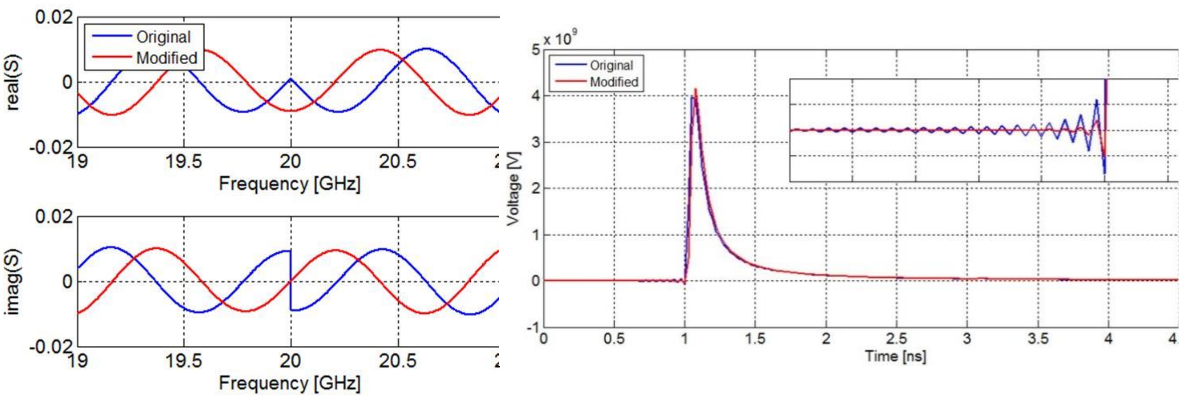


Figure 19: Comparison impulse response of S-parameters shown on Figure 18.

Limiting the bandwidth to make the S-parameters continuous changes the time domain step as well. For some applications, this is not desired. A second method to reduce the ripple is to make the imaginary part of the S-parameter zero at F_{max}. This can be accomplished by adding a small delay (between 0 and Δt) to the S-parameter. This is illustrated on Figure 20.



(a) (b)

Figure 20: S-parameters made continuous by adding a small delay. Figure (a) shows the S-parameters near F_{max}, Figure (b) shows the corresponding impulse responses. The ripple is significantly reduced.

Another method can be used to obtain causal S-parameters with a limited bandwidth that do not suffer from ringing. Assume that the continuous time domain impulse response is known. By limiting the impulse response to one time window, T_0, sampling the impulse response and performing the DFT on the sampled signal, a bandwidth limited S-parameter is obtained with a causal, time domain response. This is illustrated in Figure 21. However, one must be aware that due to spectral leakage, this bandwidth limited S-parameter differs from the continuous S-parameter, and as such, an error is introduced. This error will increase for low loss signals.

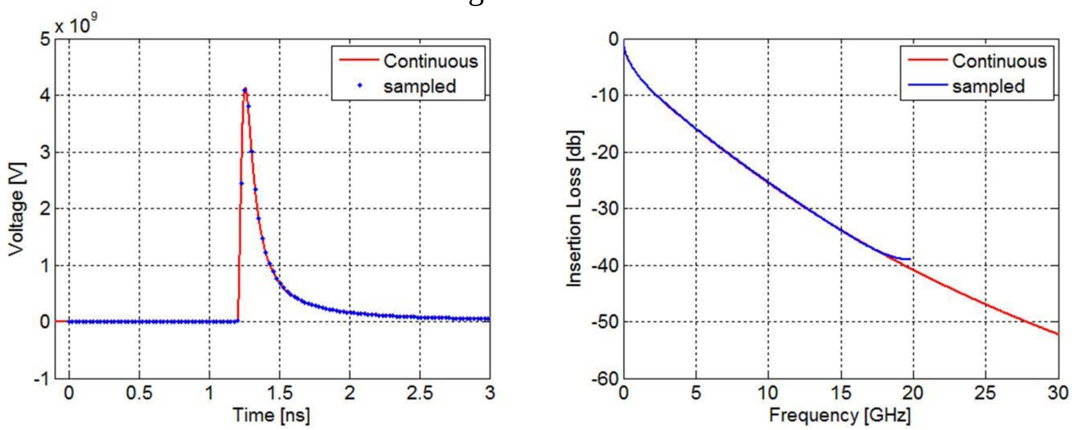


Figure 21: Continuous and sampled impulse response and corresponding Spectrum.

5. Causality for discrete bandwidth limited signals

Causality is quite simple when signals are infinite and bandwidth limited. Numerical non-causalities do not exist and physical non-causalities, if they exist, can easily be eliminated. By understanding how the DFT works and what the link is with the FT, causality definitions and enforcement can easily be transferred to discrete bandwidth limited signals.

Equation (1) for discrete bandwidth limited signals becomes

$$v_s(t_k) = 0 \quad \text{for} -\frac{T_0}{2} \leq t_k < 0 \tag{21}$$

With $v_s(t_k)$ representing the impulse response of the sampled, bandwidth limited channel and $T_0 = 1/\Delta f$ the time domain window.

Just as for the continuous signals, the causality requirement is translated into a requirement that links even and odd parts of $v_s(t_k)$. From the definition of $v_{s,even}(t_k)$ and $v_{s,odd}(t_k)$

$$\begin{aligned} v_{s,even}(t_k) &= \frac{v_s(t_k) + v_s(-t_k)}{2} \\ v_{s,odd}(t_k) &= \frac{v_s(t_k) - v_s(-t_k)}{2} \end{aligned} \qquad k = -N+1,\ldots,0,\ldots,N \tag{22}$$

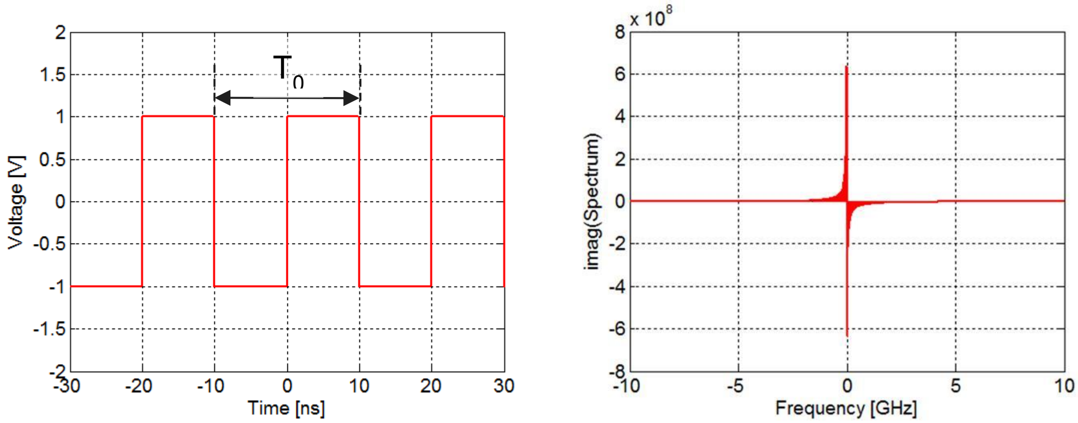
it follows

$$\begin{aligned} v_{s,even}(t_k) &= \mathrm{sign}_{T_0}(t_k).v_{s,odd}(t_k) \\ v_{s,odd}(t_k) &= \mathrm{sign}_{T_0}(t_k).v_{s,even}(t_k) \end{aligned} \qquad k = -N+1,\ldots,0,\ldots,N \tag{23}$$

with

$$\mathrm{sign}_{T_0}(t_k) = \begin{cases} -1 & -\frac{T_0}{2} < t_k + n.T_0 < 0 \\ 0 & t_k + n.T_0 = -\frac{T_0}{2},0 \\ 1 & 0 < t_k + n.T_0 < \frac{T_0}{2} \end{cases} \qquad \begin{aligned} &n = \ldots,-2,-1,0,1,2,\ldots \\ &k = -N+1,\ldots,0,\ldots,N \end{aligned} \tag{24}$$

Since the time domain signal of a sampled signal is periodic and the periodicity must be maintained, one cannot use the sign(t) function to link even and odd parts of $v_s(t_k)$, therefore one must use the periodic sign function, $\mathrm{sign}_{T_0}(t)$. An illustration of this function is shown in Figure 22.



(a) *(b)*

Figure 22: Illustration of $sign_{T_0}(t)$, (a) time domain representation, (b) frequency domain representation.

When taking the DFT of equation (23), the frequency domain representation of the causality requirement is obtained.

$$\begin{aligned} V_R(f_k) &= SIGN_{T_0}(f_k) \circledast V_I(f_k) \\ V_I(f_k) &= SIGN_{T_0}(f_k) \circledast V_R(f_k) \end{aligned} \qquad k = -N+1,\ldots,0,\ldots N \tag{25}$$

$\circledast$ is the cyclic convolution operator. This operator is needed as all functions are periodic. When a function is not causal, causality enforcement can easily be applied. Similar to the continuous case one obtains

$$v_{s,causal}(t_k) = 0.5\left(v_s(t_k) + v_s(t_k).sign_{T_0}(t_k)\right) \tag{26}$$

Or, in the frequency domain

$$V_{causal}(f_k) = 0.5\left(V(f_k) + SIGN_{T_0}(f_k) \circledast V(f_k)\right) \tag{27}$$

One must recognize that real world S-parameters will never completely meet causality conditions, thus initiating the need for a number which can indicate when causality correction is warranted.

Identical to the continuous case, the energy of the signal in the "negative time" is compared to the total energy. The smaller this number, the more causal the signal is. This causality number can be calculated in the time and frequency domains.

In the time domain, one obtains:

$$\text{Causality Number} = 100\frac{\sum_{k=-N+1}^{N}\left(v_s(t_k) - v_{s,causal}(t_k)\right)^2}{\sum_{k=-N+1}^{N}\left(v_s(t_k)\right)^2} = 100\frac{\sum_{k=-N+1}^{-1}\left(v_s(t_k)\right)^2}{\sum_{k=-N+1}^{N}\left(v_s(t_k)\right)^2} \tag{28}$$

The equivalent number calculated in the frequency domain is

$$\text{Causality Number} = 100.\frac{\sum_{k=-N+1}^{N}\left|V(f_k) - V_{causal}(f_k)\right|^2}{\sum_{k=-N+1}^{N}\left|V(f_k)\right|^2} = 100.\frac{\sum_{k=-N+1}^{N}\left|\Delta V(f_k)\right|^2}{\sum_{k=-N+1}^{N}\left|V(f_k)\right|^2} \qquad (29)$$

The previous section indicates that high quality S-parameters require that the impulse response falls completely in half the time window. The following are a few special cases in which this does not apply:

- case 1: the duration impulse of the response is larger than $T_0/2$ but smaller than T_0
- case 2: the impulse duration is larger than T_0
- case 3: the impulse duration is smaller than T_0 but the impulse response starts at t1 with $T_0/2<t1< T_0$.

The impulse response of case 1 is illustrated in Figure 23. It is clear that the impulse response is not causal due to the limited time window. Causality, however, can easily be restored since time domain leakage does not occur. It is sufficient to append the non-causal part of the impulse response to the causal part. Also notice that non-physical non-causalities, if they exist, will interfere with the numerical non-causalities and can, in general, not be removed without errors.

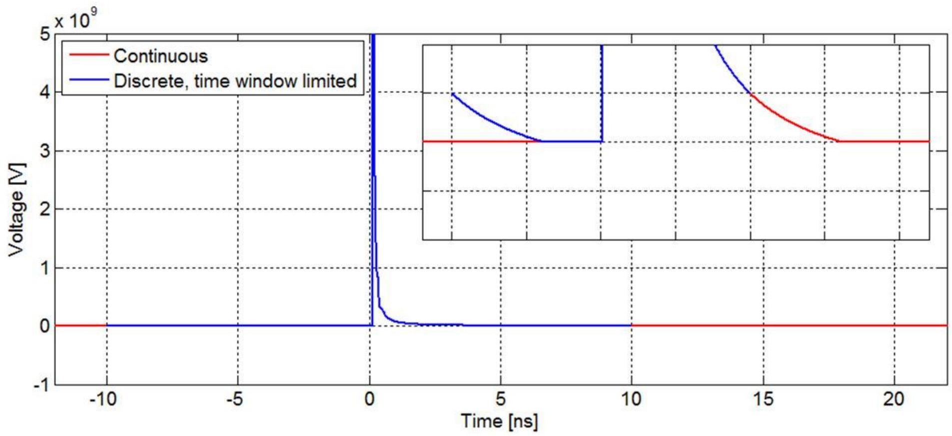


Figure 23: Impulse response with duration smaller than the time window but larger than halve the time window.

If the time window of case 1 is reduced by increasing the frequency step, Δf, then one obtains case 2. It is clear that the impulse response is not causal, and that, due to time domain leakage, the original impulse response cannot be recovered without errors.

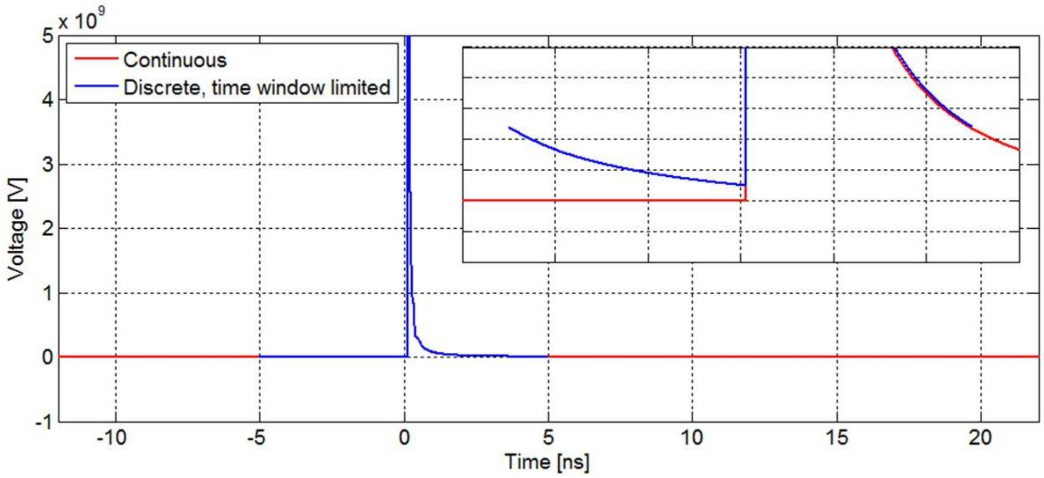
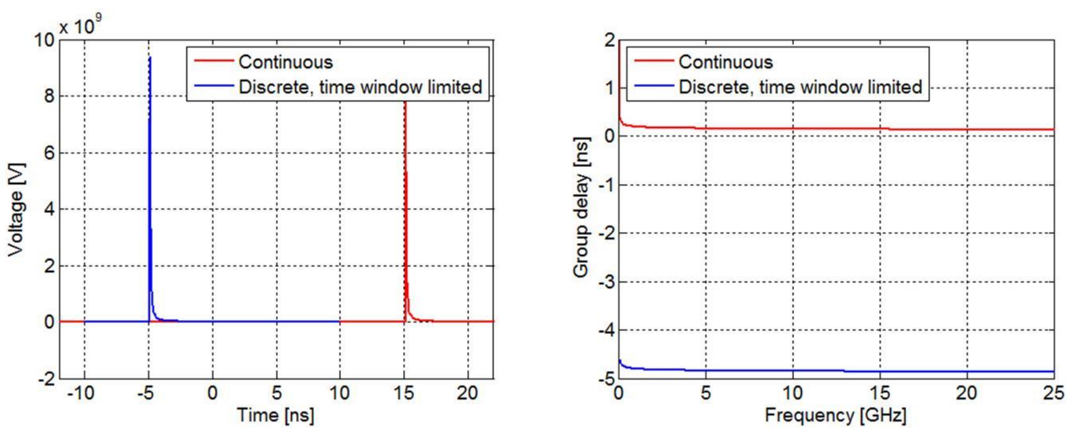


Figure 24: Impulse response with duration larger than the time window. Due to time domain leakage the tail of the impulse response interferes with the start of the impulse response and cannot be restored without making some assumptions.

Case 3 is shown in Figure 25. This case is included to show that, according to the math, the time window must be taken from $–T_0/2$ to $T_0/2$ and not from 0 to T_0. The figure shows an impulse response starting at 15 ns. Although the impulse response is smaller than T_0, which is 20 ns, the group delay is negative, and as such, considered to be non-causal. Causality can however easily be restored and causal S-parameters can be calculated. The discrete impulse response needs to be shifted over one period, $T_{0,}$ and the time window needs to be increased from $–T_0$ to T_0 by zero padding in the time domain. As a consequence, the frequency step will be reduced with a factor 2.



(a) *(b)*

Figure 25: (a) Example of impulse response that starts between $T_0/2$ and T_0, (b) Corresponding group delay.

6. Non-physical non-causalities

Along with numerical non-causalities, S-parameters can also suffer from non-physical non-causalities. These non-causalities could also be present if the frequency of the models is continuous and the bandwidth is infinite. As these non-causalities are non-physical, they must be avoided and should not be present in good quality S-parameters. Non-physical non-causalities appear frequently in S-parameters derived from measurements. Measurement noise and measurement post-processing, such as test board de-embedding and fixture removal, are important sources of non-causalities. Also calibration methodologies, such as TRL, introduce non-causalities if the launches of the calibration standards are not identical. S-parameters derived from simulations can also suffer from non-physical non causalities. Bad meshing, non-physical material properties and numerical approximations are typical sources.

Figure 26 shows an example of a non-physical non-causality. The figure shows the impedance of a connector and footprint measured with a VNA using TRL calibration. Measurements were performed with a frequency step of 25 MHz up to 20 GHz. As the results are clearly not causal, causality enforcement is required to obtain good quality S-parameters. In the first step, the ringing will be minimized by making the S-parameters, which are not continuous at F_{max}, continuous by adding a small delay to the S-parameters (Figure 27). Next, the impulse response is made causal using equation (26); the results of these steps are shown in Figure 28. Although equation (26) generates perfectly causal S-parameters, the equation also has some side effects such as changing the DC performance. To avoid this, the causality enforcement should not be applied to the impulse response, but instead to the step response. As the step response is the integral of the impulse response (equation 30), causality enforcement on the step response guarantees that the impulse response will also be causal. Once the step response is made causal, the impulse response can easily be calculated as the differential of the step response.

$$\mathrm{step}(t_k) = \sum_{k=-N+1}^{k} h_s(t_k) \tag{30}$$

$$h_s(t_k) = \begin{cases} \mathrm{step}(t_k) & k = -N+1 \\ \mathrm{step}(t_k) - \mathrm{step}(t_{k-1}) & k = -N+2,\ldots,N \end{cases} \tag{31}$$

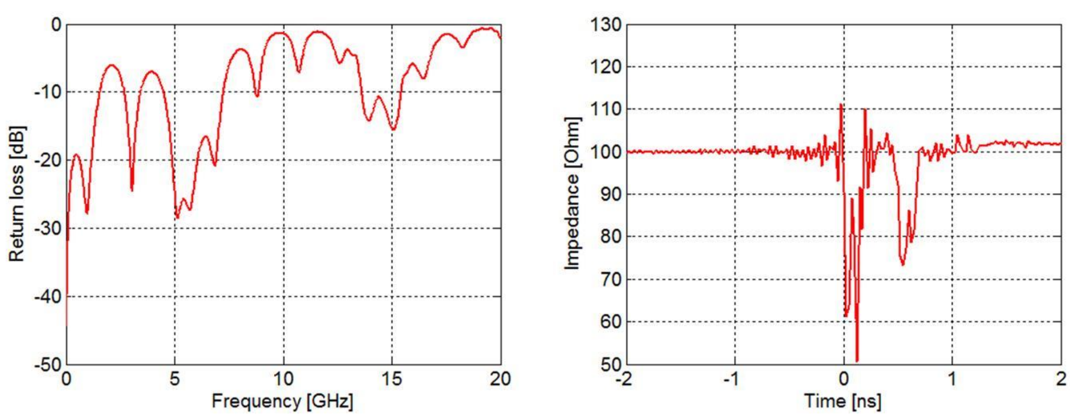
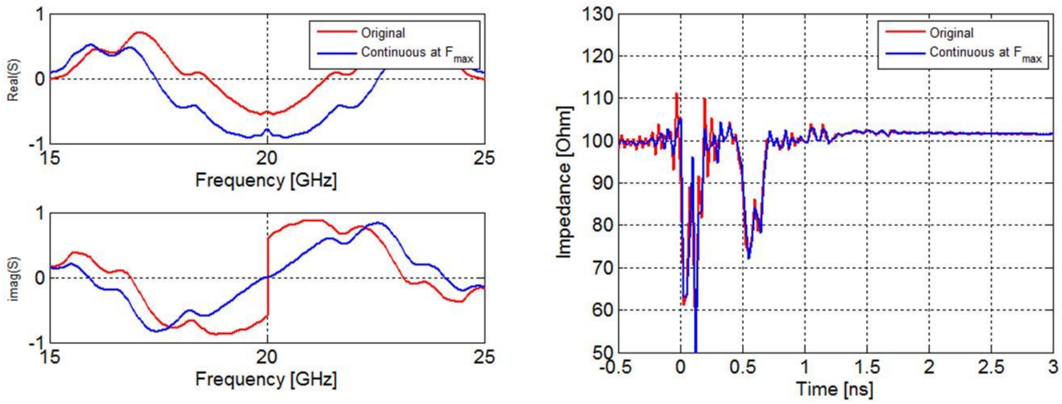
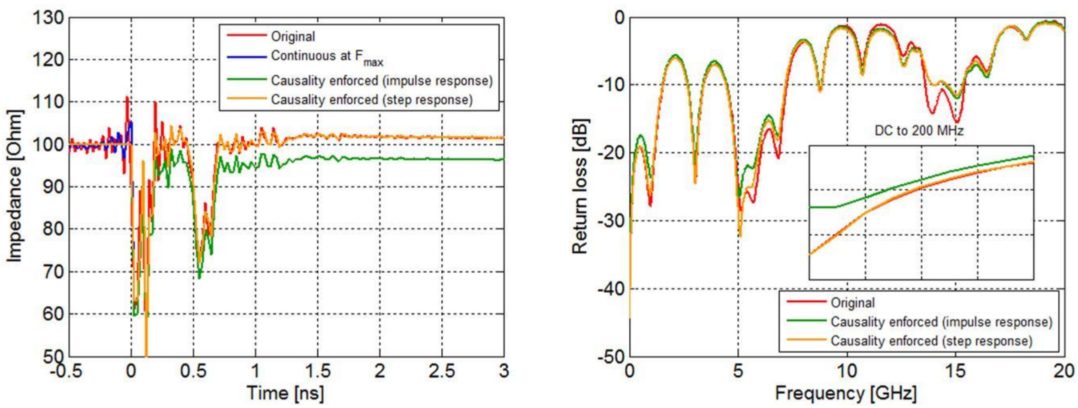


Figure 26: Example of non-physical non-causalities due the calibration errors. (a) Return loss, (b) measured impedance of a connector and footprint.



(a) *(b)*

Figure 27:(a) Real and imaginary part of the measured return loss before and after made periodic, (b) Corresponding impedance profile.



(a) (b)

Figure 28: (a) Impedance profile obtained by enforcing causality on the step response rather than the impulse response, (b) corresponding S-parameters.

7. Conclusions

This paper dealt specifically with causality for bandwidth limited, sampled S-parameters. It was shown that causality is straight forward for continuous S-parameters with infinite bandwidth, but that problems can occur if the S-parameters are sampled and limited in bandwidth.

Two different types of non-causalities exist: numerical and non-physical. Numerical non-causalities are due to the bandwidth limitations and sampling of the S-parameters. These non-causalities do not exist if the S-parameters are continuous and not bandwidth limited. Bandwidth limitations cause ringing and can only be avoided if all energy of the signal is within the considered bandwidth. Ringing can be minimized by making sure that the group delay at the maximum frequency is an integer multiple of the discrete time step.

Sampling the S-parameters can cause time domain leakage which is interpreted as non-causality. This will occur if the frequency sampling rate is smaller than the inverse of two times the impulse response duration. Causality enforcement can be applied to numerical non-causalities, but will in general introduce unwanted errors in the S-parameters. Enforcing causality on numerical non-causalities should be avoided.

Non-physical non-causalities are not related to the sampling and bandwidth limitation of the S-parameters, but are caused by measurement or simulation errors. Non-physical non-causalities interfere with numerical non-causalities, and it is not possible to separate both. Non-physical non-causalities can only be accurately corrected if there are no numerical non-causalities.

References

[1] J.S. Toll, "Causality and the Dispersion Relation: Logical Foundations", Physical Review, Vol. 104, pp. 1760-1770 (1956).

[2] R. de L. Kronig, "On the theory of the dispersion of X-rays", J. Opt. Soc. Am., Vol. 12, pp. 547-557 (1926).

[3] C. Warwick, "Understanding the Kramers-Kronig Relation Using A pictorial Proof", White paper, Agilent technologies, URL: "*cp.literature.agilent.com/litweb/pdf/5990-5266EN.pdf*"

[4] J. Bechhoefer, "Kramers-Kronig, Bode and the meaning of zero", American Journal of Physics, Vol. 79, Issue 10, pp. 1053-1059 (2011).