NEGATIVE GROUP DELAY AND CAUSALITY

Đordević, A.R. Tošić, D.V. Sch. of Electr. Eng., Univ. of Belgrade, Belgrade, Serbia

Abstract available at: http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5733913

This paper appears in: Circuits and Systems for Communications (ECCSC), 2010 5th European Conference on
Issue Date: 23-25 Nov. 2010
On page(s): 286 - 289
Location: Belgrade
Print ISBN: 978-1-61284-400-8
INSPEC Accession Number: 11875776
Date of Current Version: 17 March 2011

Negative Group Delay and Causality

Antonije R. Đorđević and Dejan V. Tošić

***Abstract*—After being passed many times from hand to hand over decades, various ideas tend to deviate from the scientific truth. Phase and group delay, and the corresponding velocities, are notorious examples. Twisting their meaning has led to some controversies. Hence, revisiting their definitions and better founding in undergraduate courses seem to be necessary.**

***Index Terms*—Group delay, Phase delay, Causality**

I. Introduction

LINEAR systems are often characterized in terms of the delay and, in the case of electromagnetic-field structures, in terms of the associated velocity. Various delays have been defined in the literature: the phase delay, group delay, signal propagation delay, and energy propagation delay [1]. However, practically (and unfortunately) only the phase delay and the group delay have remained in the academic use.

The idea and interpretations of the phase delay are properly handled throughout most textbooks and research literature. Improper meanings have been attributed primarily to the group delay, leading to confusion. This includes the existence of the negative group delay and the superluminal group velocity. These controversies are addressed by revisiting the definition and interpretation of the group delay.

II. Phase Delay and Group Delay

We consider a linear passive time-invariant physical system (a lumped-element electrical circuit, RF/microwave device, antenna system, etc.) without initially stored energy. We focus on the frequency response, $\underline{H}(\mathrm{j}\omega)$, which may represent various functions of the system (input impedance, scattering parameters, etc.). The frequency response can be represented in terms of its modulus (magnitude response) and argument (phase response) as $\underline{H}(\mathrm{j}\omega) = A(\omega)\exp(\mathrm{j}\phi(\omega))$, where $\omega = 2\pi f$ (in s^{-1}) is the angular frequency. Assuming the sinusoidal steady state, if the input signal is sinusoidal, $u_1(t) = U_1\sqrt{2}\cos(\omega t + \theta_1)$, the output signal is also sinusoidal, $u_2(t) = U_2\sqrt{2}\cos(\omega t + \theta_2)$. The corresponding complex rms signals are $\underline{U}_1 = U_1\exp(\mathrm{j}\theta_1)$ and $\underline{U}_2 = U_2\exp(\mathrm{j}\theta_2)$, respectively.

By the definition of the frequency response, the complex signals are mutually related as $\underline{U}_2 = \underline{H}(\mathrm{j}\omega)\underline{U}_1$. The rms values of the input and output signals are related by $U_2 = U_1 A(\omega)$, and the arguments by $\theta_2 = \theta_1 + \phi(\omega)$. The phase response, $\phi(\omega)$, shows for how much the phase of the input signal is to be increased to match the phase of the output signal. Hence, $-\phi(\omega)$ is referred to as the phase delay.

The phase response can be positive or negative. A phase advance does not imply noncausality of the system. A sinusoidal signal is a fictitious (non-physical) signal, which extends in time from $-\infty$ to $+\infty$. The values of this signal are predictable at any time instant. Hence, such a signal carries no information.[1]

The group delay is defined as

$$\tau = -\frac{\mathrm{d}\phi(\omega)}{\mathrm{d}\omega}, \tag{1}$$

i.e., it is the negative of the slope of the phase response. The classical way of developing the idea of the group delay is to consider a sinusoidal carrier, $u_\mathrm{c}(t) = U\sqrt{2}\cos\omega t$, whose amplitude is modulated by a sinusoidal signal, $u_\mathrm{m}(t) = \cos\Omega t$, where $\Omega << \omega$,

$$u_1(t) = U\sqrt{2}\cos\omega t\cos\Omega t$$
$$= \frac{\sqrt{2}}{2}U\big(\cos(\omega+\Omega)t + \cos(\omega-\Omega)t\big). \tag{2}$$

The spectrum of this modulated signal consists of two closely spaced spectral lines. If this narrowband signal is inputted to the system, the system output is

$$u_2(t) = \frac{\sqrt{2}}{2}U\big(A(\omega+\Omega)\cos\big((\omega+\Omega)t + \phi(\omega+\Omega)\big)$$
$$+A(\omega-\Omega)\cos\big((\omega-\Omega)t + \phi(\omega-\Omega)\big)\big). \tag{3}$$

Assuming that the magnitude response does not vary rapidly in the vicinity of ω, i.e., $A(\omega-\Omega) \approx A(\omega) \approx A(\omega+\Omega)$, and that the variations of the phase response can be approximated by the linear term in the Taylor series, i.e.,

This work was supported in part by the Serbian Ministry of Science and Technological Development under grant TR 11021 and by the COST action IC0603.

Authors are with the School of Electrical Engineering, University of Belgrade, P.O. Box 35-54, 11120 Belgrade, Serbia (phone: + 381 11 3218 329; e-mail: edjordja@etf.rs).

[1] In a less rigorous, practical view, a sinusoidal signal is a signal which becomes steady-state after all transients have passed. During the transients, the phase of the output signal adjusts itself to the steady-state value.

$\phi(\omega + \Delta\omega) \approx \phi(\omega) + \frac{d\phi}{d\omega}\Delta\omega$, we have

$$u_2(t) \approx \sqrt{2}UA(\omega)\cos(\omega t + \phi(\omega))\cos\left(\Omega\left(t + \frac{d\phi}{d\omega}\right)\right). \tag{4}$$

The term $\tau = -\frac{d\phi}{d\omega}$ is the time delay of the modulating signal, i.e., the time delay of the envelope of modulated signal. Since we consider a "packet" of two closely spaced spectral lines, i.e., a "group" of frequency components, this delay is referred to as the group delay.

Another derivation is often considered in textbooks [2] in terms of the Fourier integral, again considering a narrowband signal. This procedure is equivalent to considering the limiting case of the above presented derivation when $\Omega \to 0$. Note that the spectrum of the modulated signal is limited. Hence, the modulated signal must be of infinite duration in time, i.e., it exists for $-\infty < t < +\infty$.

The modulating signal does not carry any information, as it is a sinusoidal signal, and therefore fully predictable. The group delay is, hence, just the time delay of the sinusoidal modulating signal (i.e., it is proportional to the phase delay of the envelope) in the steady state!

For electromagnetic-field systems (transmission lines, waveguides, etc.), where we have to consider signal propagation along certain physical distances, the phase delay and the group delay are associated with the corresponding velocities: the phase and the group velocity, respectively.

The concept of various delays and velocities has been considered for a long time. Sommerfeld and Brillouin [1] considered wave propagation in dispersive materials. They identified several different velocities: phase, group, energy, signal propagation, the first forerunner, and the second forerunner. The phase and group velocities are associated with the phase and group delay as defined above for a steady-state response. The energy velocity is somewhat artificially defined in terms of the density of the electromagnetic-field energy and the Poynting vector of the wave. The first forerunner is the signal precursor, i.e., the very beginning of the signal at the output, it is followed by the second forerunner, and then the bulk signal occurs. Physically, the velocities of the forerunners and the signal velocity cannot exceed the speed of light in a vacuum (Einstein's causality). A good survey of these concepts can be found, for example, in [3].

However, some researchers and textbook writers have been neglecting the differences among these delays and velocities, which can create misinterpretation of the underlying physics. They have been presenting primarily the phase delay and group delay (and the corresponding velocities), messing up their properties in addition. Perhaps the worst misconception was published in [4], where the phase velocity was identified with the information-propagation velocity and it was hence concluded that the information can be propagated faster than light.

As a more subtle example, the group velocity is often identified with the velocity of propagation of energy or information [5]. This has physical explanations in some cases. An example is a lossless metallic rectangular waveguide, where the propagating wave can be interpreted as a uniform plane wave that bounces, being reflected from the four waveguide walls. However, such an interpretation is wrong in many other cases, in particular when the system contains a standing wave or lossy elements. In such systems, the group delay can be very small, even negative, as demonstrated and clarified in [6] for various passive and active circuits. For electromagnetic-field systems, a very short (positive) group delay has been experimentally observed in some cases. This phenomenon has brought up its own controversies of superluminal velocities, which have been examined both theoretically and experimentally. Many researchers have made improper conclusions, though physically sound and rigorous explanations for the phenomenon are available in the literature [7].

In the papers that present a proper explanation for the negative group delay, only the bulk effect on the signal is considered (e.g., broadband signals), and the effect of the group delay on the envelope of a modulated signal is not examined. In [8], the group delay is defined as the time delay imposed on the envelope of the signal as a result of its passing through a channel. A constant modulus of the channel frequency response across the signal bandwidth is assumed. Following this classical book, the negative group delay was mixed with the issue of frequency variations of the magnitude response. The "invisibility" of the noncausal response was improperly attributed to the interference between the magnitude and phase responses.

Our intention is to clarify that **the group delay can be negative** and how the negative group delay affects the narrowband-signal propagation.

To illustrate the existence of the negative group delay, we consider several physically realizable passive electrical networks. In the analysis of lumped-element circuits, the physical dimensions of the system are assumed vanishing. Hence, the output signal can appear at the same instant as the input signal. For transmission lines and other electromagnetic-field systems, the signal has to travel a certain physical distance, so that the output signal always starts delayed after the input signal.

The first network, shown in Fig. 1, is a lossless transmission line, which is mismatched at both ends. The input reflection coefficient of the line is shown on the Smith chart in Fig. 2. (Note that the reflection coefficient can be mapped into the transmission coefficient of a two-port network, without any additional delay, by using a resistive directional coupler.) It is a periodic function of frequency. As the frequency increases, the argument (phase response) of the reflection coefficient decreases and increases periodically. The group delay alternates the sign and is positive and negative, as shown in Fig. 3.

The impulse response of this system cannot be easily computed numerically, as the spectrum is infinite and periodic.

However, it can be computed in the time domain, using, e.g., a SPICE-based circuit solver. The impulse response consists of a train of delta-functions. The first delta-function occurs at zero time. The delay between adjacent impulses is 5 ns, which is two times the signal propagation down the line. The amplitudes of the impulses decay fast. No dispersion is present, and the response is causal, as should be expected from a physical system.

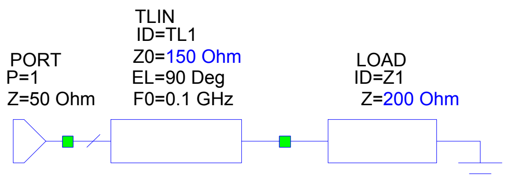

Fig. 1. A transmission line mismatched at both ends.

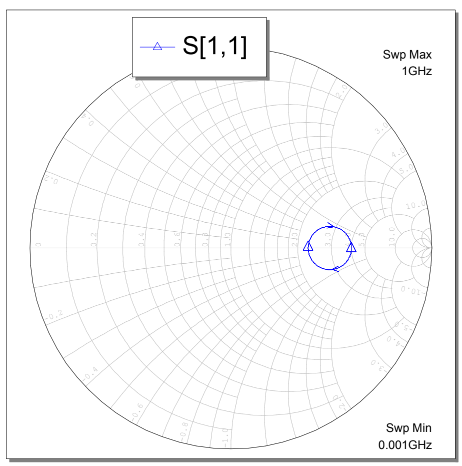

Fig. 2. Input reflection coefficient of the network shown in Fig. 1.

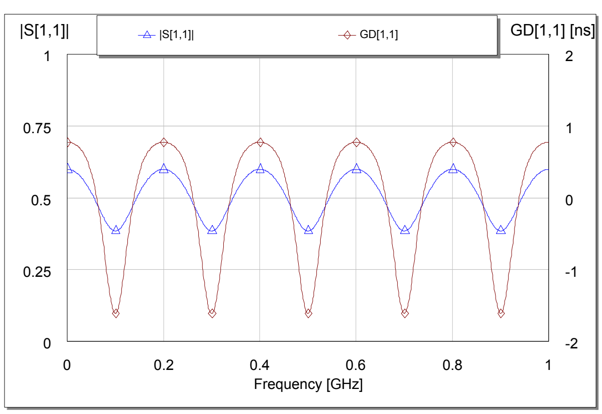

Fig. 3. Modulus of the input reflection coefficient and the group delay of the network shown in Fig. 1.

Even a simple lumped-element lossy network can exhibit a negative group delay. An example is the network shown in Fig. 4, which is the model of a real coil, up to the first resonant frequency. The reflection coefficient of this network has a frequency range where the argument increases with frequency, resulting in a negative group delay. A similar observation holds for the input admittance, shown in Fig. 5. The frequency range where the argument increases, corresponds to a negative group-delay region. The largest modulus of this negative delay is around 350 MHz and it amounts to about 2 ns.

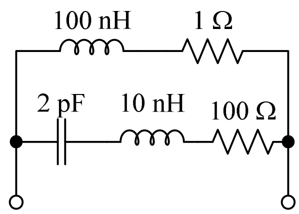

Fig. 4. Lumped-element equivalent network of a coil.

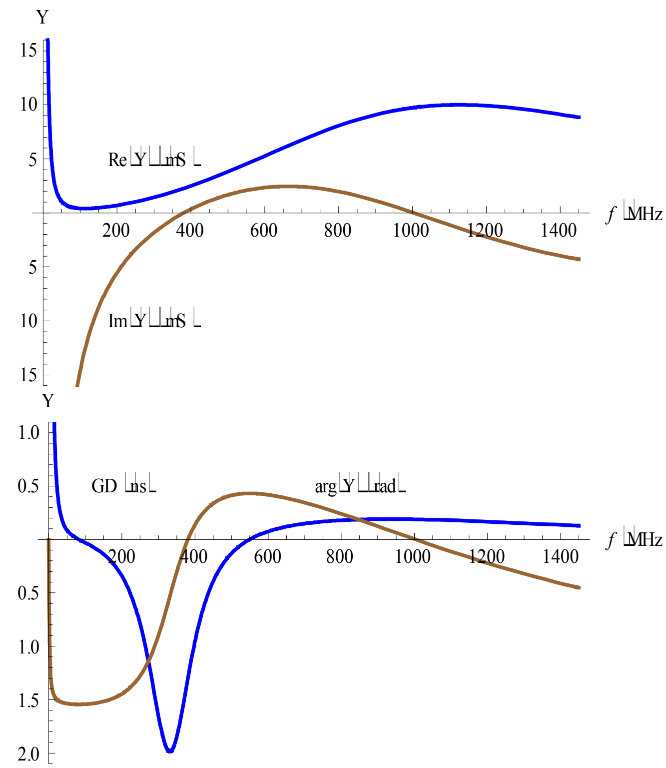

Fig. 5. Real and imaginary parts, group delay, and argument of the input admittance of the network shown in Fig. 4.

To give an interpretation of the negative group delay, we consider the network of Fig. 4 driven by a voltage generator, and observe the current at the terminals. Its electromotive force is a sinusoidally amplitude-modulated carrier of frequency 350 MHz. The frequency of the modulating signal is 5 MHz. The modulated signal is multiplied by a rectangular pulse, whose leading edge is located at $t = 0.1\ \mu\text{s}$, so that the resulting electromotive force is zero for $t < 0.1\ \mu\text{s}$. The width of the pulse is $T = 0.72\ \mu\text{s}$. This allows to clearly mark the starting point of the excitation, in order to be able to investigate the causality. The driving electromotive force is shown in Fig. 6. In the frequency domain, the corresponding spectrum (Fig. 7) consists of two peaks. However, the peaks are not sharp due to the modulation by the rectangular pulse.

The response of the system is the current at the terminals and it is causal, as shown in Fig. 8. Fig. 9 shows a zoom-in of the input and output waveforms. The output waveform is causal and it starts synchronously with the excitation. The envelope of the modulated signal starts from zero, so that we can consider it to be exactly in phase with the input signal. After some time, however, the envelope of the output signal attains a phase advance with respect to the envelope of the input signal. This corresponds to the negative group delay, and it is clearly visible by comparing the two signals in the vicinity of time instants where the envelope diminishes to zero (Fig. 9).

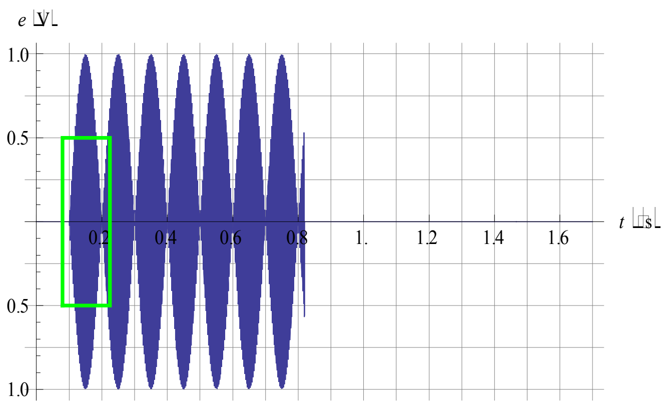

Fig. 6. Electromotive force of the generator driving the network shown in Fig. 4.

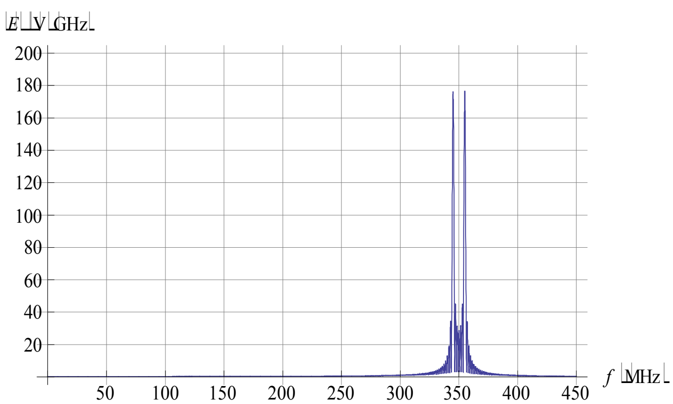

Fig. 7. Amplitude spectrum of the electromotive force shown in Fig. 6.

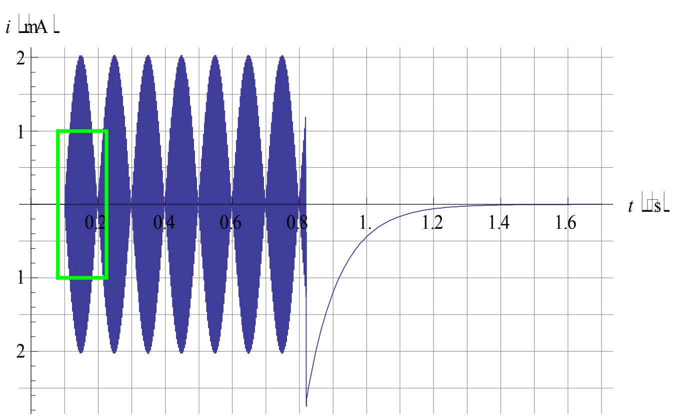

Fig. 8. Current of the network shown in Fig. 4.

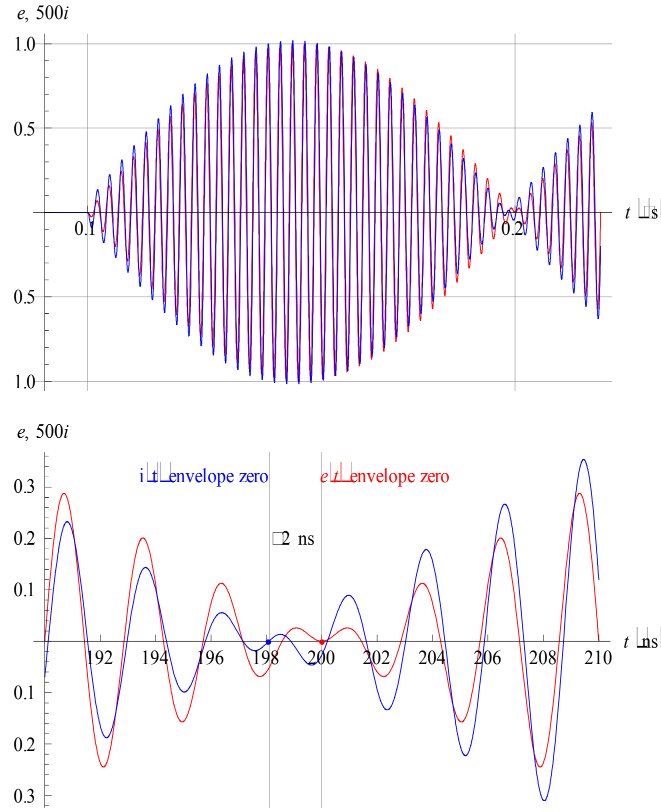

Fig. 9. Zoom-in of the waveforms shown in Fig. 6 and Fig. 8.

III. Conclusion

The aim of this paper is to revisit some frequently misinterpreted facts about the phase delay and group delay, as well as the group and phase velocities. We have demonstrated that the group delay for physically realizable systems can be negative. Hence a negative group delay, does not imply noncausality. The results presented here might provide background for teaching undergraduate courses.

References

[1] L. Brillouin, *Wave propagation and group velocity*, Academic Press, New York, 1960.

[2] I. S. Stojanović, *Fundamentals of Telecommunications*, BIGZ, Belgrade, 1972.

[3] K. Zhang and D. Li, *Electromagnetic Theory for Microwaves and Optoelectronics*, Springer, Berlin, 2007.

[4] T. K. Ishii and G. C. Giakos, "Transmit radio messages faster than light", *Microwaves and RF*, vol. 30, pp. 114-119, 1991.

[5] B. D. Popović, *Electromagnetics*, Gradjevinska knjiga, Belgrade, 1980.

[6] M. W. Mitchell and R. Y. Chiao, "Causality and negative group delays in a simple bandpass-amplifier", *Am. J. Phys.*, vol. 66, pp. 14-19, 1998.

[7] M. Mojahedi, E. Schamiloglu, K. Agi, and K. J. Malloy, "Frequency domain detection of superluminal group velocities in a distributed Bragg reflector", *IEEE J. Quantum Electron.*, vol. 36, pp. 418-424, 2000.

[8] A. Papoulis, *Signal Analysis*, McGraw-Hill, New York, 1977.