Vladimir V. Kopeikin

Wave refraction in linear media with time dispersion

NAUKA

Vladimir V. Kopeikin

Wave refraction in linear media with time dispersion

MOSCOW NAUKA 2009

UDC 530.1

Reviewers:
A.P. KISELEV, Dr. Sc. (Phys. & Math.),
V.P. VOLCHKOV, Dr. Sc. (Tech.)

Kopeikin Vladimir V.
Wave refraction in linear media with time dispersion / Vladimir V. Kopeikin. – Moscow : Nauka, 2009. – 128 p. – ISBN 978-5-02-036962-7 (in cloth).

The monograph proposes a new method for deriving spatial-temporal geometric optic equations based on more general equations of wave energy transfer in the space than the equations used in a standard method. A new approach makes it possible to pass from a locally plane homogeneous monochromatic field model, which is an exact solution to the wave equation for a homogeneous medium and forms the basis for the standard version of geometric optics, to a smoothly inhomogeneous model in the form of the Airy function, which is an exact solution of the wave equation for the problem of refraction in a linear layer. The obtained approximation describes the effect of dispersion refraction and corrects the description of usual refraction for unmodulated monochromatic waves in linear media with time dispersion.

This monograph can be used by specialists in the field of wave processes, by radiophysics studying radio propagation in the ionosphere, and by students and post-graduates of the corresponding specialities.

ISBN 978-5-02-036962-7

Foreword

Any moving source of waves excites a wave process with transverse frequency modulation in the ambient space due to the Doppler Effect. Wave frequency is maximal or minimal along or against the direction of source motion, respectively. This frequency gradually varies from maximum to minimum between 0 and π. Transverse frequency modulation is maximal at an angle of $\pi/2$ [13].

Any wave packet (finite in time and space), propagating in an inhomogeneous medium with time dispersion, is subjected to transverse frequency modulation because the spectral components of this packet diverge at different angles; i.e., the longitudinal spectrum is transformed into the transverse spectrum [62].

This indicates that transverse frequency modulation is far from exotic but is the most common phenomenon in physics of wave processes, especially taking into account the fact that all inhomogeneous media, strictly speaking, have a time dispersion to a greater or lesser extent.

Only stationary point or monochromatic sources in homogeneous media cause waves without transverse frequency modulation.

Until recently, it has been considered that transverse frequency modulation does not affect wave refraction in dispersive media. This is at least confirmed by the standard ray optic equations. This monograph indicates that this is not the case and it is necessary to review the approaches to description of refractive phenomena in linear media with time dispersion.

The refraction concept is almost always related to the concept of a ray optic or ray approximation. The refraction phenomenon can be characterized as a change in wave propagation trajectory as a result of wave interaction with a medium, described within the scope of ray optics (RO). In an RO approximation, a certain quantity n called refractive index or optical density is introduced. Sometimes, refractive index can be explicitly expressed in terms of wave field and physical parameters of a medium (electron concentration, pressure, density, wave amplitude, etc.) and is specified in the original wave equation (from which an RO approximation is derived) as a phenomenological parameter [1–4, 6, 8, 11, 16, 69].

The aforesaid does not indicate that only RO can describe refraction; however, spatial and time scales of refraction phenomena characterized by an insignificant change in the wave trajectory in the space and time interval of fast field oscillations are determined in this approximation.

The spatial-temporal scales of the refraction effects are unambiguously related to similar scales of a change in refractive index; i.e., the n value should also be "smooth" on the wavelength and "slow" in the time interval of fast field oscillations. Otherwise, one should speak about diffraction wave effects [5, 10, 12, 18, 35, 39, 40, 50, 64].

Linear and nonlinear refractions are usually considered. If refractive index depends only on characteristics of a propagation medium, linear refraction is considered. If refractive index depends not only on these parameters but also on wave intensity, nonlinear refraction is considered.

Both refraction types are described by their own ray approximations: ray approximation itself and nonlinear ray approximation. At present, these ray approximations are applied to the spatial-temporal case, and linear and nonlinear space-time ray optics (STRO) is considered. Nevertheless, these approximations have a common feature: the wave process is represented in the form of a locally plane homogeneous and monochromatic wave, the parameters of which "smoothly" and "slowly" change during propagation. Some researchers used similar representations in the form of, e.g., non-monochromatic functions, but we will ignore these representations because they are well-known and are within the scope of the linear and nonlinear refraction conceptions [9, 11, 31, 36, 46, 58].

Moreover, the generalization of both stationary and space-time ray methods is intensely developed for the case of locally inhomogeneous plane waves, the so-called RO of inhomogeneous waves or complex RO. In this approximation the solution is sought in the form of a plane wave with a complex wave vector **k**, the imaginary component of which describes transverse field attenuation and is commensurable with the real component responsible for wave length. Therefore, the term "ray optics" is, generally speaking, not quite correct because this approximation also takes into account diffraction effects. These effects result in that a complex ray in a homogenous medium is generally not a straight line [12, 18, 41–45, 56, 59–60, 63].

However, one less known refraction mechanism is associated with the medium time dispersion properties and is the main subject of the present study.

We should note once more that all material media interacting with a wave to a certain extent have time dispersion, which manifests itself in the dependence of refractive index n on frequency [1–4, 17, 26–30, 38, 40, 51, 79]. The dependence of refractive index n on field intensity results in a special class of wave phenomena: nonlinear refraction; similarly, the dependence of n on frequency results in the other class of wave phenomena: hereafter, dispersive refraction (DR).

Dispersive and nonlinear refraction have a number of common properties because refractive index depends on wave field in both cases. For example, DR can manifest itself in a homogenous medium, which is impossible in the case of ordinary linear refraction. A difference between dispersive and nonlinear refraction consists in that nonlinear refraction can originate when amplitude is modulated along a wave front, whereas DR (in the case of a quasi-monochromatic wave packet) appears only when frequency is modulated along this front. In a certain sense, DR can be considered among nonlinear effects because it is determined by the nonlinear relationship of frequency ω and wavenumber k in the dispersion equation. Consideration of the dispersive properties of a medium results not only in the propagation features that can be related to the DR effect and whose scale is much larger than the wavelength but also in the features that can be called dispersive diffraction and the scale of which is commensurable with the wavelength. Such features (in particular, the effect of plasma "blooming" by an inhomogeneous wave) will be discussed in Chapter 1; however, the monograph mainly studies the refractive phenomena.

The following, physically nonrigorous, qualitative considerations will explain the DR nature.

Assume that refractive index n is the function of frequency, and a quasi-monochromatic wave packet, with an oscillation frequency changing along the wave front, propagates in a homogenous dispersive medium. Figure 1 shows the wave function corresponding to the ultra-broadband wave packet with a transverse frequency modulation.

The phase velocity in different front zones is different owing to the dispersion relation; therefore, according to the Huygens principle, the wave front behavior should evidently be the same as in a dispersion-free but inhomogeneous medium with a spatially variable refractive index: the front propagation direction should change; i.e., the refraction effect should be observed. From the viewpoint of ray optics, a medium affects a wave via refractive index, and the cause of spatial variations in n is of no fundamental importance for the wave: this cause can be either inhomogeneity of medium properties, as in the case of linear refraction, or frequency inhomogeneity, as in our case. At nonlinear and dispersive refraction, the medium can have homogeneous physical parameters, and refractive index inhomogeneity can depend on a spatially inhomogeneous wave-medium interaction [4–6, 26–30, 74, 93].

We now briefly consider the development of the theoretical concepts of wave processes in dispersive media and try to analyze why the DR effect is little-known.

A classical theory of wave dispersion in linear media was completed by Zommerfeld and Brilluen [1] at the beginning of the last century. They examined all known wave effects associated with wave dispersion, developed a mathematical apparatus used to describe these effects, and designated the general way for solving all linear non-stationary problems, based on the signal expansion into the Fourier spectrum and on the superposition principle. Subsequently, the interest to the non-stationary problems decreased because it became evident that any such problem can be solved when the stationary field is known in some frequency range. Researchers mainly tried to find the methods for integrating stationary wave equations since these equations are simpler than non-stationary ones. At that time, all works devoted to dispersion considered, as before, one-dimensional or quasi-one-dimensional problems, when the DR effect could not be detected in principle.

The Rytov's dissertation [31] published in 1940, which formed the basis for STRO, played an important role in the development of the mathematical methods in the theory of wave propagation in dispersive media. However, STRO started intensely developing only in the 1960s. In contrast to the classical approach, the stationary problem is not solved, and an approximate solution of the non-stationary wave equation is immediately sought in STRO.

The usage of STRO resulted in widening of the scope of solvable problems because this method made it possible to start studying non-stationary wave processes in complex two- and three-dimensional media close to real ones [3, 11, 16].

In a classical approach, it is very difficult to perform such studies because methods that would make it possible to exactly solve stationary wave equations in inhomogeneous multidimensional media are still absent. Computational approaches can hardly be applied to such media even in the stationary problem, which is part of the non-stationary problem, because the possibilities of computers are limited.

A standard STRO version of the space-time ray method, which is based on the field representation in the form of a locally plane homogenous monochromatic wave, certainly does not describe diffractive effects comparable with a wavelength. However, this version also does not describe the DR effect related to a week transverse frequency inhomogeneity of the wave field.

A standard field model is an exact solution of the wave equation for a homogenous medium. Such a "plane" model arises when only two main terms of the infinite ray series (eikonal and transfer equations) are taken into account and the contribution of the rejected terms is ignored. However, if we consider the applicability of the standard field model more attentively, it turns out that the cumulative contribution of the rejected components in a dispersive medium cannot be discarded because the value of this contribution could be comparable with that of the refraction effects.

If we take the rejected components into account and use the local field expression in the form of the Airy function (which is the exact solution of the wave equation in a linear layer) as a model, we will obtain a more rigorous STRO version, which not only describes the DR effect but also corrects the description of ordinary linear refraction. The monograph will indicate that the STRO standard version is, strictly speaking, applicable only to media without time dispersion and is only partially applicable to dispersive media (concerning general RO conditions (4.1)).

Nevertheless, until recently the STRO standard version has been successfully used to describe wave propagation in dispersive media (mainly HF radio propagation in the Earth's ionosphere), because the DR effect in a real medium is insignificant for ordinary narrowband signals emitted by usual antennas.

The correction, which arises when monochromatic wave refraction is described strictly, becomes significant at near-critical frequencies, and this range is usually of no interest at a "hoping" radio propagation in the ionosphere. The error of traditional ray optic is often large, but this disagreement between the theory and experiment is usually explained by the fact that ionospheric parameters are unknown during observations.

Recently, the situation has changed because broadband and even ultra-broadband signals, the relative bandwidth of which approaches unity, have been used along with narrowband radio signals [80, 81, 101]. In real media the DR effect is substantial for such signals and sometimes defines the wave energy propagation mechanism. In such cases the STRO standard version cannot be used.

Below we will briefly describe the history of the radio engineering development from the viewpoint of applied signal types.

Until the mid-1970s, almost all radio systems used the following signals to transmit information:

$$s(t) = A(t)\sin[\omega_0 t + \psi(t)]. \tag{F.1}$$

These signals are usually called quasi-monochromatic, quasi-harmonic or sinusoidal signals. Here $A(t)$ and $\psi(t)$ are relatively "slow" modulating functions compared to the fast oscillations defined by the $\sin\omega_0 t$ function. Various methods of signal modulation (in amplitude, frequency, phase, etc.) indicate that these signals are narrowband; i.e., the spectral energy is concentrated in the vicinity of a carrier frequency ω_0. The relative signal bandwidth $\Delta\omega/\omega_0$, where $\Delta\omega$ is the bandwidth, generally does not exceed 1% and is not larger than 0.1%, e.g., in the HF range.

Ground penetrating radars (GPRs), which operate in the ultra-broad band without a carrier frequency, appeared in the late 1960s. It turned out that only such signals, sometimes called videopulses, make it possible to perform subsurface sounding. The point is that the soil conductivity is very high, as a result of which radiofrequency is strongly absorbed. It is necessary to decrease a carrier pulse frequency in order to increase sounding depth and to reduce the envelope length in order to provide required depth resolution. Videoimpulse is the limit for such trends. Thus, the present-day radio engineering uses almost the same signal types that were used by G. Hertz, A.S. Popov, and G. Markoni during the first experimental electromagnetic energy transmissions over large distances [76, 85, 87–91, 94].

At the beginning of the last century, a passage from ultra-broadband signals, emitted by spark transmitters, to quasi-harmonic signals was caused by the necessity to increase the communication distance and to prevent the communication systems from interference of adjacent radio stations. The problem was solved using the oscillating circuit as a selective unit for quasi-harmonic signals. Extraordinary simplicity of the oscillating circuit, which consists of a capacitor and inductance coil, was responsible for future development of radio engineering. Mechanical quartz-crystal resonators, coaxial line sections, strip lines, and box units, were subsequently used along with regular oscillating circuits. However, all of them are selective units for type (F.1) signals.

Thus, an exclusive usage of quasi-monochromatic signals was explained by purely technical causes because radiation and radio propagation theory does not fundamentally restrict a signal bandwidth. For example, any specified physically realizable wave field function is transmitted in a dispersion-free non-absorbing medium without distortions.

Until the end of the 1940s, a rigorous mathematical theory of radio reception was absent, and hardware engineers mostly used experimental approaches. A number of scientists (e.g., K. Shennon, N. Wiener, A.J. Khinchin) developed a theory of optimal reception based on a probability theory and mathematical statistics. V.A. Kotel'nikov [84] considerably contributed to the optimal receiving theory, which confirmed that the oscillating circuit is an optimal filter for quasi-harmonic signals [80, 81].

According to the theory, an optimal receiver consists of a correlator that performs convolution of a received signal $x(t)$ with a model signal $s(t)$.

$$y(t) = \int_{-\infty}^{\infty} x(t-\tau)s(\tau)d\tau. \tag{F.2}$$

For signals of (F.1) type this procedure is performed by an oscillating circuit. The type of the $s(t)$ function is not restricted in an optimal receiving theory. For example, in the problem of signal detection against a background of additive normal white noise, receiver specification depends only on a noise spectral density and signal energy, and is independent of signal waveform and, consequently, on a spectrum bandwidth.

Recently, the development of computer engineering has made it possible to use signals different from (F.1). In advanced radio receivers, convolution (F.2) is performed by a microprocessor, which can find the $y(t)$ function for an arbitrary $s(t)$ signal, and the advantages of quasi-monochromatic signals are lost in this case.

At present, a new signal class becomes actual: broadband and ultra-broadband noise-like signals (NLSs). This signal provides maximum possible information transmission rate as well as stealthiness and noise protection close to the theoretical limits. Code Division Multiple Access (CDMA) or Ultra Wide Band Code Division Multiple Access (UWB CDMA) is realized in this case.

The spectral density of NLSs is uniform within the operating frequency band, and the level of this density is lower than that of natural noise, as a result of which the system with NLSs is not fixed by standard radio equipment. Therefore, communication is absolutely physically stealthy (i.e., an object of potential cryptographic deciphering is absent in this case).

Noise-like radio systems can operate against a background of ordinary narrowband stations, almost not affecting their operations, and vice versa: narrowband stations almost do not affect noise-like station operation. The number of noise-like stations that can operate in the same frequency band is much larger than that of ordinary stations, other things being equal. This is of great importance for the development of radio engineering because the frequency resources are already limited.

Technical problems of creating communication systems with NLSs (from LF to UHF signals) are principally resolvable, and intense works have been performed in this field [80, 86, 89, 92]. However, when developing such systems, one should face the problem that is absent (or not so acute) in quasi-harmonic systems: the effect of a dispersive medium on broadband or ultra-broadband radiosignals.

This problem is especially critical in the HF range. During the ionospheric propagation of a signal within the band wider than 100 kHz, the DR effect causes not only a dispersion distortion of a signal waveform but also a change in the propagation mechanism. At the same time, we will subsequently indicate that broadband signals make it possible to control radio propagation in the ionosphere by multiply increasing the communication energy potential, stealthiness, and noise protection as well as the information rate.

The aforesaid indicates that the monograph subject-matter, including the usage of the monograph results (NLSs), is of current importance. This monograph describes not only broadband radio propagation; the DR effect can be observed in the wave processes of any kind in any dispersive media, including a homogeneous medium and narrowband signals.

The DR effect originates in the presence of transverse frequency modulation of waves. Such inhomogeneous waves in a homogeneous medium can be generated by physically realizable emitters, such as a distributed source, a point source system, or a moving transmitter (see above). Only special emitters can be used to generate a packet with transverse frequency modulation in a homogeneous medium. Similar emitters can generate transverse field inhomogeneity for both narrowband and broadband signals.

As was noted above, the DR effect always exists in an inhomogeneous dispersive medium if a wave is non-monochromatic. For wave packets emitted by ordinary antennas, a longitudinal spectrum is transformed into a transverse spectrum in an inhomogeneous dispersing medium, after which transverse frequency modulation and the DR effect originate. This effect depends on the structure of wave packet and space inhomogeneity. For example, if special emitters are not used for the ionosphere with its characteristic parameters, the DR effect becomes substantial when a signal band is wider than 100 kHz.

The main aim of this monograph is to analyze the DR effect from different viewpoints and to obtain the main result: a modified version of the equations for space-time RO, which take this effect into account.

The monograph consists of seven chapters.

The first chapter includes the original wave equations, which describe wave propagation in a media with time dispersion. The general integral-differential equation makes it possible to describe waves in a medium with an arbitrary dispersion function. Special attention is paid to the hyperbolic Klein–Gordon equation (KGE), which specifies the plasma-like dispersion function. This is the most common dispersion type in the HF radio band since this dispersion exists in the ionosphere, which is the largest Earth's sphere. A number of explicit solutions, which describe the physics of wave processes in dispersive media (e.g., circular wave motion in a homogeneous space), are presented for this equation.

Several examples of the KGE explicit solutions in inhomogeneous media, which indicate that a plane homogeneous wave model cannot rigorously describe refraction in a dispersive medium, are also presented here.

The second chapter is devoted to wave packet propagation in a homogeneous half-space with and without time dispersion. The explicit solutions for the wave packet integral characteristics (mean group velocity vector, transverse coordinate, propagation time, width, and duration) were obtained in the general form for arbitrary boundary conditions. It has been shown, that wave packet motion in a homogeneous medium is not uniform and rectilinear at certain boundary conditions. Sufficient conditions were formulated for the existence of the DR effect. The wave field functions for the packet with transverse frequency modulation in homogeneous dispersive and dispersion-free media were numerically calculated based on the Green function. A transverse shift of packets (the DR effect) in a dispersive medium is evidently traced in the presented figures.

The third chapter begins with the consideration of the necessary condition of DR effect origination in a homogeneous medium: the ways of emission of wave packets with transverse frequency modulation. A moving monochromatic point source is the simplest prototype of such an emitter. In this case transverse frequency modulation originates due to the Doppler Effect.

This chapter is mainly devoted to the non-stationary parabolic equation, which can be used to describe DR. This equation can be used in numerical computations because the complex wave amplitude smoothly varies along coordinates, which makes it possible to effectively use the finite-difference scheme with a large step. However, along the longitudinal coordinate, the dispersion equation differs from the original KGE; therefore, the parabolic approximation applicability conditions are presented. An example of the numerical computation is presented for the wave packet with transverse frequency modulation in dispersive and dispersion-free homogeneous media. The DR effect is pronounced in a dispersive medium.

The fourth chapter considers a new method for deriving equations of space-time RO. This method makes it possible to use more complicate field models than a locally plane homogeneous monochromatic wave, which is used in standard STRO. It has been demonstrated that the standard model has an error in the linear component of the original KGE approximation, which corresponds to the error on the scale of refractive effects.

The plane field model originates when the ray series is described only by the eikonal transfer equations without numerous amplitude corrections. However, these corrections cannot be rejected in the case of a dispersive medium because the infinite sum of corrections does not vanish in the refraction problems.

The approximation error was eliminated, and the modified equations of space-time RO with the DR effect were obtained by correcting the model (by substituting a smoothly inhomogeneous model). Moreover, ordinary refraction was decreased using corrective terms.

An approach to the derivation of the ray equations, proposed in the previous chapter, in the case of arbitrary time dispersion is developed in the fifth chapter. The equations, taking into account DR and correcting the description of ordinary refraction, are also obtained in this chapter. A quasi-ray field model, which makes it possible to describe a change in the global field parameters (e.g., transformation of longitudinal frequency modulation into transverse modulation), is introduced.

In the sixth chapter, the DR effect is considered based on the principle of monochromatic wave superposition. In this respect, DR is a two-dimensional (or three-dimensional) version of space-time focusing. The possibility of using DR to artificially control radio propagation in the ionosphere is considered. It has been demonstrated that this method makes it possible to increase the output performance of HF radio systems by several orders of magnitude.

The systematic error of a standard STRO, originating when propagation of non-modulated quasi-monochromatic waves in the ionosphere is described, is analyzed in the seventh chapter. The DR effect is absent for such waves, but an additional term, correcting ordinary linear refraction, is present in the modified version of ray optics.

The value of the systematic error is the problem of great practical importance because the standard version of ray optics is widely used to analyze radio propagation in the ionosphere. Both STRO versions (standard and modified) are compared to the exact wave solution, and the causes of the systematic error are discussed in the first section of this chapter.

The results of numerical ray computations, obtained using two methods for certain ionospheric models, are presented and compared in the second section of this chapter. It is demonstrated that the error can be intolerably large in some cases, e.g., when radio waves emitted from a satellite are calculated. The error is almost imperceptible in other problems, such as description of a hoping propagation of radio waves in the ionosphere.

In conclusion the equations of a modified STRO version are presented, and it is explained why a smoothly inhomogeneous field model should be used to obtain ray equations in a dispersive medium in contrast to a dispersion-free medium.

I am grateful to Mr. A.V. Popov for repeated useful discussions of the monograph subject-matter.