Formation of Ionospheric Precursors of Earthquakes—Probable Mechanism and Its Substantiation

Georgii Lizunov¹, Tatiana Skorokhod¹, Masashi Hayakawa², Valery Korepanov³

¹Space Research Institute, Kyiv, Ukraine
²Hayakawa Institute of Seismo Electromagnetics Co., Ltd., Tokyo, Japan
³Lviv Center of Institute for Space Research, Lviv, Ukraine

Email: georgii.lizunov@gmail.com, hayakawa@hi-seismo-em.jp, vakor@isr.lviv.u

Abstract

The purpose of this article is to attract the attention of the scientific community to atmospheric gravity waves (GWs) as the most likely mechanism for the transfer of energy from the surface layers of the atmosphere to space heights and describe the channel of seismic-ionospheric relations formed in this way. The article begins with a description and critical comparison of several basic mechanisms of action on the ionosphere from below: the propagation of electromagnetic radiation; the closure of the atmospheric currents through the ionosphere; the penetration of waves throughout the neutral atmosphere. A further part of the article is devoted to the analysis of theoretical and experimental information relating to the actual GWs. Simple analytical expressions are written that allow one to calculate the parameters of GWs in specific experimental situations. Specificity of GW dispersion properties and features of their propagation are analyzed on this mathematical basis, processes of amplitude amplification and dissipation of GWs with height are investigated, the mechanism of generation of ionosphere-magnetosphere current systems is described and their quantitative characteristics are determined. The experimental part presents an analysis of GWs global distribution in the thermosphere derived from the data of the instrument NACS (Neutral Atmosphere Composition Spectrometer) onboard the satellite DE-2 (NASA, 1981-1983). The statistical association of registered ionospheric disturbances with earthquakes is demonstrated. The results of DE-2 data processing are backed up by comparison with data from the DEMETER satellite (CNES, 2005-2010) whose purpose was to study the ionospheric effects of earthquakes. Specific features of GWs that characterize these waves as a factor of influence on the ionosphere from below are indicated.

DOI: 10.4236/ojer.2020.92009   Mar. 16, 2020   142
Keywords

Atmospheric Gravity Waves, Seismic-Ionospheric Coupling, Thermosphere

1. Introduction

Recently, the concept of earthquake (EQ) preparation has been formed as an ensemble of mechanical, chemical and electromagnetic phenomena, initiated by tectonic processes and propagating upward into the atmosphere and the ionosphere (e.g., [1] [2]). The geophysical anomaly preceding the EQ is formed as a large-scale 3D structure based on the source of a future EQ and stretching into outer space. The study of such a structure requires approaches and techniques far beyond of classical seismology scope.

The holistic picture of the seismic-ionospheric interaction may be divided into several relatively independent blocks: 1) processes in the lithosphere and the lower atmosphere preceding the EQ and generating surface anomaly of geophysical parameters (emanation of core gases, local changes in surface temperature, etc.); 2) propagation of the anomaly through the atmosphere to space heights by means of various mechanisms of energy transport in the atmosphere-ionosphere-magnetosphere system; 3) generation of wide range of geospace plasma disturbances, taken as ionospheric precursors of an EQ (electromagnetic fields, plasma inhomogeneities, sporadic layers of the ionosphere, variations in the fluxes of energetic particles, etc.). This work is devoted to the study of the second and partly of the third block. Its main goal is to re-draw attention to the internal atmospheric gravity waves (hereinafter—GWs) as the most promising agent of the seismic-ionospheric interaction.

In our study, we try to bring together theoretical and experimental information about the properties of GWs penetrating to the ionosphere from below. As a theoretical basis, we give the approximate expressions that facilitate calculation of the GW parameters in specific experimental situations. Using them the readers can easily reproduce all the calculations of this work. Moreover, the use of rigorous mathematical theory of GWs for the analysis of experimental data or theoretical modeling is often an excess of calculation accuracy. This is due to the fact that the parameters of the atmosphere in the altitude range from the earth’s surface to the ionosphere are never exactly known (and, in turn, are being taken from models of the atmosphere), as well as the fact that the basic theory of GWs does not take into account a number of really valid factors (dissipation, nonlinearity, inhomogeneity of the atmosphere, etc.).

Our study is based on the key publications in this area [3]-[10]. But these papers do not cover a number of topics important for understanding the ionospheric effects of GWs, such as their amplification and dissipation with altitude, electromagnetic response generation of the ionosphere, experimental GW characteristics obtained from observations on low earth orbit satellites. Our aim is
to fill these gaps.

2. Mechanisms of Energy Transport from Surface Sources to the Ionosphere

The most discussed mechanisms for the energy transport from surface sources to the ionosphere are: 1) low frequency electromagnetic radiation; 2) quasi-static electric fields and currents; 3) waves of the neutral atmosphere. Following Rishbeth [11], we note that any ionospheric disturbances, including those caused by EQ preparation processes, obey the law of matter conservation: ion concentration change = ionization – recombination +/– transport. Thus, the effectiveness of the mechanisms of impact on the ionosphere from below depends on their ability to influence photochemical reactions and plasma transport processes. From this point of view the mentioned mechanisms efficiencies will be compared below.

This work is mainly devoted to the study of effects at altitudes of the E- and F-regions of the ionosphere, accepted to be above 100 km.

**Electromagnetic radiation.** Electromagnetic radiations (EMR) caused by surface sources are observed from space with the launch of first satellites. Moreover, EMR was the first described type of ionospheric EQ precursors (according to measurements on the Intercosmos-19 satellite [12] [13]).

Energetically, the main near-Earth sources of EMR are lightning discharges. The worldwide thunderstorm activity determines the features of the electromagnetic environment of our planet, first of all—the stable structure of Schumann resonances with central frequencies 7.8 Hz, 14.3 Hz, 20.8 Hz, 27.3 Hz, etc [14] [15]. Next are whistler waves, whose frequencies range from units to several tens of kHz and which serve as lightning discharge markers observable from satellites. In addition, numerous EMR of technogenic origin are recorded in the same frequency bands: power lines emissions and their harmonics, signals of navigation transmitters and broadcasting stations, etc., whose spectra show signs of nonlinear and resonant interactions with the ionosphere plasma [15]-[20].

If we exclude the active experiments of heating the ionosphere with powerful radio emission, it can be argued that the energy of natural and technogenic electromagnetic radiation is, as a rule, insufficient to modify the regular parameters of the ionosphere (Although electromagnetic pulses from lightning discharges cause detectable heating of the electron component; its localization region is limited by the heights of the D-layer) [21].

**Quasi-stationary electric field.** The fundamentals of the atmospheric electricity theory were developed in the middle of the twentieth century and are exposed, for example, in [22]. The initial equations of quasi-stationary electromagnetic field are,

\[ \nabla \cdot \mathbf{j} = i, \quad \mathbf{j} = \sigma \cdot \mathbf{E}, \quad \mathbf{E} = -\nabla \phi, \]

where \( \mathbf{j} \) is the electric current density, \( i \) is the density of current sources, \( \sigma \) is the electrical conductivity, \( \mathbf{E} \) is electric field, \( \phi \) is electric potential; outside of
sources \( \nabla \cdot j = 0 \). Any conclusions regarding quasi-stationary effects in the ionosphere should not contradict the solution of these equations.

A common misconception is already the idea of the earth-ionosphere gap as a capacitor in which the lower plate (Earth’s surface) is negatively charged, the upper plate (ionosphere) is positive, and the plates are separated by a quasi-neutral atmosphere. Ionosphere perturbations in this model are interpreted as a result of changes of the capacitor plates charge. In fact, there is no upper plate: the atmospheric gap between the Earth and the ionosphere is positively charged by itself. Moreover, the positive space charge is in direct contact with the Earth’s surface, and the altitude level with a potential of 250 kV, above which the atmosphere is to be considered an almost equipotential conductor, is located only at the foot of the ionosphere, at an altitude of about 60 km. So, the entire configuration of the Earth electric field is enclosed within the troposphere and the stratosphere, forming some kind of the “Faraday cage”, which shields the ionosphere from electrical interference from below [22].

As is known, after the electromotive forces action termination, the space charge in a conducting medium disappears in time \( \tau \sim \varepsilon_0 / \sigma \) (where \( \varepsilon_0 \) is vacuum permittivity). This time is about 5 minutes on the Earth’s surface; at altitude of 40 km (at the lower boundary of the D-region) it is about 0.1 s. At higher altitudes, particle collisions become so rare that electron inertia should be taken into account when calculating conductivity; in this case, the estimation of charge neutralization time has the form \( \tau \sim \sigma / \left( \varepsilon_0 \omega_p^2 \right) \), where \( \omega_p \) is plasma (Langmuir) frequency. In the E-region, \( \tau \sim 10 \) ms. At larger times, charge separation should be maintained. Thus, an uncompensated electric charge in the atmosphere and the associated electric field arise as a result of the stationary circulation of electric currents generated by sources of atmospheric electricity and partially closed through the Earth below and the ionosphere above [22] [23]. So, the concept of an Earth-ionosphere capacitor (D-region) only makes sense when considering EMR with frequencies of tens of Hz or more.

As an example, Figure 1 shows the electrical circuit configuration of a thundercloud. The cloud charge model is represented as an elementary dipole. Then

![Figure 1. Electric field of a thundercloud. The current source is represented as an elementary dipole.](image_url)
the electric field above the earth is the sum of the dipole field and its electrostatic image at the Earth’s surface; inside the earth it is determined by the field of the initial dipole, depressed proportionally to the ratio of earth and air conductivity, i.e. $\sim 10^{10}$ times [22]. Because of an exponential conductivity increase with height above the cloud, a part of the electric current lines are stretched upwards to smaller resistance. With the cloud height increase the part of the current closed through the ionosphere grows [23]. The opposite is also true: as lower the current source is located, the less is upward electric flux. So, it is clear that the leakage of currents from underground magnetotelluric sources into the ionosphere is negligible. But as to the current generated at the upper layer of clouds, a significant part of it reaches the E-region, where it flows into the global ionosphere-magnetosphere electric circuit. However, the fact is that this current is extremely small $\sim 10^{-6} - 10^{-5} \mu A/m^2$ [24] [25], which is five to six orders of magnitude less than the background currents flowing in the E-region [26].

The situation is not better with the penetration of the Earth’s electric field into the ionosphere, though at the Earth’s surface it reaches significant values ($\sim 100$ V/m at fair weather and even $\sim 1$ kV/m under thunderclouds). To confirm this, we consider a vector tube formed by the current lines above, for example, the cloud in Figure 1. From the equation $\nabla \cdot j = 0$, the total current through the cross-section $S$ of the tube is not dependent on the height:

$$j_0S_0 = j(z)S(z) = \sigma(z)E(z)S(z) = \text{const},$$

where the index 0 refers to the lower edge of the tube. (It should be clarified that the parameter $\sigma$ here is a combination of the components of the atmospheric conductivity tensor, in the simplest case $\sigma = \sqrt{\left(\sigma_\parallel \cos \theta\right)^2 + \left(\sigma_p \sin \theta\right)^2}$, where $\sigma_p$ is the Pedersen conductivity, $\sigma_\parallel$ is the parallel conductivity, $\theta$ is the angle between the direction of the electric field and the Earth’s magnetic field. Neglecting the change in the tube cross-section related to the divergence or focusing of current lines with height, we obtain:

$$j = j_0 = \sigma(z)E(z) = \sigma_0E_0 = \text{const} \quad \text{or} \quad E(z) = \frac{\sigma_0}{\sigma(z)}E_0 = \frac{j_0}{\sigma(z)}.$$  

Thus, the near-surface electric field perturbations are transmitted upwards with the attenuation factor $\sigma_0/\sigma(z)$. The vertical distribution of the electrical conductivity tensor components of the earth-atmosphere-ionosphere system is presented in Figure 2. The effective electrical conductivity $\sigma(z)$ increases from the ground to the level of 60 km (base of the D-region) by 5 - 8 orders of magnitude, and to the level of 120 km (E-region)—by 10 - 13 orders of magnitude (depending on the mutual direction of the electric and magnetic fields along current line). Respectively, attenuation factors of the electric field are $10^{-5} - 10^{-4}$ and $10^{-10} - 10^{-13}$. A reasonable estimation of the field of surface sources at E-region is about 0.1 μV/m, which is three orders of magnitude less than background ionospheric fields at the same heights.

Contrary to the foregoing, a number of publications devoted to the numerical simulation of the penetration of surface electric fields into the ionosphere claim...
that much larger field values (up to 1 mV/m in the $E$ region) have to be generated there. Such overestimates, apparently, are the result of an error in the formulation of the upper boundary condition for the differential problem (1). This question was investigated in detail in [27], where numerical experiment has been posed correctly. Figure 3 shows an example of calculating the vertical distribution of the electric field of a model surface source 100 V/m from this work. The plots on the graph correspond to different heliophysical conditions.

**Neutral atmosphere waves.** From the theoretical hydrodynamics point of view, the atmosphere of the Earth (as well as atmospheres of any stars and planets) is a gravitationally stratified medium. Because of this atmosphere waves have the properties strongly different from those of electromagnetic waves, due to which they may be considered as a powerful energy transport mechanism upward from the inner layers of the atmosphere to the outside. The key role here is the exponential decrease in the atmosphere density with altitude—the barometric distribution, due to which even the energetically weak variations of the near-surface air layers may spread to ionospheric heights, where they become strong in the scale of the rarefied space environment. In such a way the increase of the ratio of wave energy density to the environment energy density between the earth surface...
**Figure 3.** Numerical calculating the altitude distribution of the horizontal component of electric field $E_x$ produced by surface source above a point with the horizontal coordinate $x_m$, at which the maximum field is reached (adapted from [27]). The horizontal size of source is taken to be 100 km, the vertical electric field on the surface $E_z = 100 \text{ V/m}$. The upper split of the plot corresponds to different integral conductivities of the ionosphere.

and a height of, say, 120 km reaches $\sim 10^7$. We note here an opposite analogy with the properties of the electrostatic channel when affecting the ionosphere from below.

Another fundamental consequence of the density barometric distribution is the growth with height of the coefficients of kinematic viscosity and thermal conductivity of the atmosphere, inversely proportional to the gas density. As a result, atmospheric waves dissipate in a peculiar way: the wave energy is absorbed in a certain altitudinal layer. Its location does not depend on the initial wave amplitude and is determined by wave spectral parameters—frequency, wavenumber and propagation direction. In such a way atmospheric waves not only propagate through the ionosphere, inducing the dynamic response of its plasma component, but also are absorbed in a certain altitudinal layer changing the temperature regime and photochemical equilibrium there.

This allows us to consider atmospheric waves as the most efficient channel for the energy transport into the ionosphere from below [28] [29]. In a wide range of atmosphere waves we select the gravity mode, propagation of which is associated with the buoyancy forces action—heated gas floats up, while cooled gas sinks. And tectonic processes just create the thermal sources necessary for the generation of such waves. According to [2], a list of phenomena occurring over the source of a forthcoming EQ includes: the emanation into the atmosphere of core gases (including greenhouse components CO$_2$, CH$_4$, etc.) and radioactive
radon; the condensation of atmospheric moisture molecules ionized by radon with an intensive release of latent heat; the mixing and convection of heated air layers. By this, a stationary dynamic structure is formed—a flow of thermal energy from ground level to the top of the clouds, which is apparently the missing link that connects atmospheric oscillations with tectonic processes.

To confirm this postulate, let us remind that seismogenic thermal anomalies are permanently recorded by remote sensing satellites in the far-infrared range in the pending EQ area. The typical parameters of the anomalies are as follows [30] [31] the heated zone corresponds to the projection of the EQ area on the Earth’s surface, reaching hundreds of kilometers in size; temperature difference with the surrounding region may be a few degrees; the lifetime of the anomaly is from several days to several weeks. Theoretical estimation of the generated atmospheric oscillations by such heat sources are presented in early works [31] [32] [33] and recent work [34]; relevant experimental evidence is given in [35] [36].

Let us consider in more details main physical processes involved in formation and propagation of the atmospheric GWs.


Atmospheric waves. The neutral atmosphere wave spectrum includes three modes depending on frequency: internal gravity waves—GWs, relatively high-frequency acoustic waves and low-frequency planetary waves. In this list, GWs play a dominant role both in their prevalence (the frequency of registration) and their energy. The fact is that the gravity mode is, on one hand, not too fast, and on the other hand, not too slow. Horizontal phase velocities of GWs can be arbitrarily small, therefore GWs are excited by numerous sources: thermal anomalies of the Earth’s surface, movement of weather fronts, atmospheric convection, etc. (in opposite to acoustic waves which are mostly generated by supersonic sources). On the other hand, GW group velocities are tens and hundreds of meters per second, which is quite a lot. Propagation time of the GW front between the Earth’s surface and the F-region of the ionosphere may be less than an hour. (Unlike planetary waves, whose propagation to the ionosphere takes weeks and months. Moreover, due to the extremely low phase speed of planetary waves, their propagation is effectively blocked by zonal wind systems in the mesosphere).

GW dispersion law. The dispersion equation of GWs [3] is not too complicated, but is quite cumbersome for visual perception. For analytic calculations, we recommend using the approximate formulas [10] [37] [38]. At any rate, the upper atmosphere parameters are usually known with low accuracy that rather devalues the application of an exact dispersion law to the analysis of experimental data.

GW circular frequency:

\[ \omega = \frac{c k_s}{\sqrt{1 + \left(\frac{k}{k_g}\right)^2}} = \omega_g \sqrt{\frac{k_s^2}{k^2 + k_s^2}}. \]
GW attenuation decrement:

\[ \nu = D\left(k_z^2 + k_g^2\right). \]  (4)

Here \( k = \{k_x, 0, k_z\} \) is wave vector. The OZ axis of the Cartesian coordinate system is directed vertically upwards, OX axis is selected in such a way that the wave vector lies in the XOZ plane;

\[ \omega_g = \sqrt{\left(1 - \gamma^{-1}\right)g/H + (dT/dz)(g/T)} \]  is Brunt-Väisälä frequency, which is the maximum frequency of the GW;

\[ k_g = \left(2H\right)^{-1} \]  is a typical scale for GW wavenumbers;

\[ c_g = \omega_g / k_g \]  is the maximum GW horizontal phase velocity, approximately equal to the speed of sound;

\[ H = RT/\mu_g \]  is the atmospheric scale-height;

\( D \) is the kinematic viscosity coefficient of the atmosphere;

\( R, \gamma, \mu, T, g \) are gas constant, adiabatic index, mean molecular weight, atmospheric temperature, and gravitational acceleration. The values of some of these parameters are given in Table 1. On the phase plane, the GW dispersion curves fill the region shown in Figure 4.

From (3) the horizontal phase velocity of GW is

\[ \omega / k_x = c_g / \sqrt{1 + (k / k_g)^2}. \]  

According to this expression \( \omega / k_x \) monotonously decreases with increasing wavenumber and is always less than the speed of sound (\( \omega / k_x < c_g \)). This property of GW should be recognized as exceptional, because the projections of a phase velocity are usually not subject to restrictions from above.

Using (3) we also obtain the expressions for components and direction of the group velocity:

\[ V_{gx} = c_g \frac{1 + \left(k_z / k_g\right)^2}{\left(1 + \left(k / k_g\right)^2\right)^{3/2}}; \quad V_{gz} = c_g \frac{k_z}{\left(1 + \left(k / k_g\right)^2\right)^{3/2}}; \]

\[ \tan \psi = \frac{V_g}{V_{gz}} = -\frac{k_z}{k_g^2 + k_z^2}. \]  (5)

The minus sign in the second and third expressions means that the vertical components of the group and phase velocities of the GW are directed oppositely (if \( k_z < 0 \), then the energy propagates upward).

In Figure 5, the vertical component of the group velocity is represented as a function of the period and of the horizontal wavelength. The distribution has the

<table>
<thead>
<tr>
<th>Table 1. Parameters used for the GW theory development.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conditions</td>
</tr>
<tr>
<td>-----------------------</td>
</tr>
<tr>
<td>altitude 0 - 150 km</td>
</tr>
<tr>
<td>altitude &gt; 150 km</td>
</tr>
<tr>
<td>high solar activity</td>
</tr>
<tr>
<td>altitude &gt; 150 km</td>
</tr>
<tr>
<td>low solar activity</td>
</tr>
</tbody>
</table>
Figure 4. Dispersion plane: GW area (1), dissipation area (2), forbidden area (3).

Figure 5. The normalized vertical component of GW group velocity $\frac{V_g}{c_g}$ as a function of the normalized horizontal wavelength $\frac{k_x}{k_g}$ and the normalized period $\frac{\omega_g}{\omega}$.

form of a ridge, which breaks in the direction of supersonic speed $\frac{\omega}{k_x} > c_g$. In the longer wavelength region $\frac{k_x}{k_g} > 1$, the ridge line is given by the equation $\frac{\omega}{k_x} = 0.7c_g$. The vertex of the distribution is reached at the point $k_x = -k_x = k_g$, $\omega = \omega_k / \sqrt{3}$ where $V_g = V_{g \text{max}} = c_g / 3^{1/2}$, $V_g = 2c_x / 3^{1/2}$. Thus, the fastest upward wave transfer occurs obliquely, at an angle to the horizon $\psi = \arctan(1/2) \approx 30^\circ$.

We assume that in Figure 5 the crest of group velocity selects a spectral window in which the vertical transport of GW energy is most effective. The window parameters are as follows: GW periods are from ten to several tens of minutes, horizontal wavelengths from fifty to several hundred kilometers, and horizontal phase velocities of more than a hundred meters per second. These data are in
good agreement with the observed parameters of medium-scale traveling ionospheric disturbances associated with the effects from below. For the productive generation of such long-wave perturbations, their surface sources should have horizontal dimensions of the same order as the wavelengths—in this case, mesoscopic scales. Such sources are known: these are thermal anomalies over the preparation area of the strong EQs, conglomerates of thunderstorm clouds, tsunami, and weather fronts. At the same time, multiple quasi-point generators of atmospheric oscillations—a plowed field heated by the sun, the work of an industrial enterprise, the wind over urban buildings, etc.,—will not have the size to be significant factors affecting the ionosphere from below.

*Amplitude amplification and dissipation of waves with height.* The absorption of GW energy, as well as any macroscopic movements of the atmosphere, occurs under the influence of viscosity and thermal conductivity effects, which are approximately characterized by a single kinematic coefficient of viscosity/thermal conductivity/diffusion \( D = D(z) \) (Figure 6). The value of this coefficient on the Earth’s surface is quite small (~10\(^{-3}\) m\(^2\)/s), but grows with height according to the exponential law. Due to this, for the same perturbation wavelengths, the atmosphere behaves as an almost perfect liquid at low altitudes and as extremely viscous and heat-conducting fluid at high altitudes. The theory of GW dissipation during propagation from bottom to top [38] [39] [40] reveals an analogy with Chapman’s theory describing the absorption of solar EUV radiation propagated from top to bottom. The equation of the vertical GW energy transfer is:

![Figure 6](image-url)
\[
\frac{dS_z}{dz} = -2\kappa S_z = -Q,
\]
(6)

where \( S_z = V_0 \rho \tilde{V}^2 / 2 \) is the density of the vertical energy flux, \( \rho \) is the density of the atmosphere, \( \tilde{V} \) is the amplitude of the particle velocity, \( Q \) is the heating rate per unit volume of the atmosphere, \( \kappa \) is the spatial attenuation decrement of the wave (absorption coefficient), which is expressed through the temporal attenuation decrement (4) using the formula \( \kappa = \sqrt{V_0} \).

In the model of isothermal and chemically homogeneous atmosphere, the wave group velocity does not change with altitude, and the density and kinematic viscosity coefficient, and the wave attenuation decrement depends on the height according to simple exponential laws: \( \rho \sim \exp\{-z/H\}, \kappa - D \sim \exp\{+z/H\} \). In this case, Equation (6) has an analytical solution:

\[
Q = Q_m \exp\left\{1 + \frac{z-z_m}{H} - \exp\left\{\frac{z-z_m}{H}\right\}\right\},
\]
(7)

where the maximum energy absorption rate \( Q = Q_m \) is achieved at a height level \( z_m \) determined by the condition:

\[
\kappa(z)\big|_{z=z_m} = 1/(2H).
\]
(8)

It may seem that the GW absorption height should depend on wave initial amplitude—weaker waves have to decay closer to the source, stronger—at greater height. But Equation (8) denies this—the parameter \( z_m \) depends on the atmosphere height scale and, through the attenuation decrement \( \kappa \), on the spectral parameters of the wave, but not on its initial amplitude. This unexpected result confirms that even weak ground-based GW sources will transport the energy to ionospheric heights, and each monochromatic component of the wave will have its proper absorption height.

To clarify this conclusion let us analyze the GW amplitude change with height. The solution of Equation (6) regarding the particle velocity is:

\[
\tilde{V}(z) = \tilde{V}_0 \sqrt{\frac{\rho_0}{\rho(z)}} \exp\left\{-\int_0^z \kappa(z)dz\right\} = \tilde{V}_0 \exp\left\{\int_0^z \left(\frac{1}{2H} - \kappa(z)\right)dz\right\},
\]
(9)

where \( \tilde{V}_0 \) is velocity at the surface. Thus, GW evolves under the influence of two competing factors: amplification with increment \( 1/(2H) \) due to a change in the density of the atmosphere with height, and attenuation with decrement \( \kappa = \kappa(z) \) due to viscosity and thermal conductivity. At low altitudes where the dissipation is low the amplification factor prevails, at high altitudes attenuation prevails. The maximum amplitude of the GW is achieved under the condition \( \kappa = 1/(2H) \) which just coincides with (8). As an example, vertical distribution (9) is plotted in Figure 7 for a wave with the following parameters: horizontal wavelength 15 km, period 11 minutes, initial velocity \( \tilde{V}_0 = 2 \text{ cm/s} \). Note that the factor of the wave amplitude amplification at an altitude of 120 km is \( 10^3 \) times.

Figure 8 shows the configuration of the GW field of a model point source, calculated on the basis of the above theory. The atmosphere is modeled by two
Figure 7. Vertical distribution of the GW amplitude.

Figure 8. GW field generated by a model point source. The dependence of the maximal height of GW propagation is shown at a low solar activity (LSA) and a high solar activity (HSA). The numbers near the plots are horizontal wavelength in km.

layers stitched at an altitude level of 150 km with the viscosity kinematic coefficient \( D = 10^5 \text{ m}^2/\text{s} \). The layer parameters are shown in Table 1. Due to the obvious simplification of the model the numbers given in the figure should be in-
interpreted as approximate.

Thus, the propagation pattern of GWs has a characteristic funnel shape. Directly above the source the waves are absent—due to the rapid attenuation of GWs which are propagating at large angle to the horizon. The maximum heights of the wave field are 200 - 300 km (depending on solar activity) at distances of more than 1500 km from the source. It should be noted that the curves presented in the figure characterize the height of maximum energy absorption $z_m = z_m(x)$; in fact the wave energy leaks higher, within approximately still one scale-height. On the other hand, only a small part of GW spectral power—in the vicinity of group velocity ridge in Figure 5—reaches maximal heights. The majority of harmonics are absorbed below in the region of gray shaded space in Figure 8.

At the intersection of wind structures, GW ray paths are refracted in such a way that the wavenumber decreases in a tailwind, and increases in the headwind. This is easy to understand from the following considerations. In the reference frame of the wind, the wave horizontal phase velocity increases if the wind is heading, and therefore, according to the dispersion Equation (3), the wavenumber $k$ decreases. On a tailwind, the opposite happens. So, at headwind the GW attenuation decrement $\kappa \sim (k^2 + k_E^2)^{1/2}$ decreases, and this leads to the GW penetration to greater heights than shown in Figure 8. At tailwind the decrement increases, and the penetration height decreases. Due to these reasons, GWs from surface sources reach the upper ionosphere irregularly, depending on the velocity and direction of thermosphere circulation relatively to the GW propagation direction.

In upper atmosphere the wave amplitude amplification reaches gigantic values. One can expect that the wave reaches the level of overturning $\tilde{V} \sim \omega/k$ and then balances near this level, periodically releasing excess energy and momentum into the atmosphere. In this case, the ionospheric disturbance has not to depend on the power of their source below. This conclusion allows us to clarify the experimental fact found in [41] as a result of processing data from the DEMETER satellite. Namely, the amplitude of ionospheric responses to EQs was almost the same in the magnitude range $M = 4.8 - 8.0$ (see Figure 2 in their paper).

Numerous data indicate that at altitudes of 70 - 150 km GW dissipation causes the atmosphere heating by tens of Kelvin. Thus, surface thermal anomalies—the sources of GW generation—are in some way transferred to ionospheric heights. Ultimately, GWs are a significant factor in the formation of the upper atmosphere energy balance (yielding in the hierarchy of sources and sinks of a heat to the absorption of solar EUV radiation, but along with the heat influx from the thermosphere and along with the cooling due to IR radiation) [7] [10] [42].

**Electromagnetic disturbances generation.** Due to the finite electrical conductivity of the atmosphere and the presence of the Earth’s magnetic field, any atmospheric movements—wind, tides, wave disturbances—are accompanied by the electric current generation:
\[ j = \hat{\sigma} \cdot \left( E + \hat{V} \times B \right), \quad (10) \]

where \( \hat{\sigma} \) is the tensor of atmosphere/ionosphere conductivity, \( \hat{V} \) is the velocity of neutral particles, \( B \) is the Earth’s magnetic field induction, and \( E \) is the electric field strength. In this expression, the second term in brackets describes the current source created by the movement of the atmosphere (dynamo-current), and the first term is the conduction current. The fundamentals of the theory of electromagnetic disturbance generation during the GW propagation have been developed in [43] [44].

The vector product \( \hat{V} \times B \) in Equation (10) indicates that the excitation of the dynamo-current occurs due to the atmosphere movement transverse to the magnetic field. Also the dynamo-current parallel to the transverse motion component occurs under the action of “double transverse” Hall conductivity. Let’s explain this on the assumption that the Earth’s magnetic field is directed vertically along the OZ axis, the wave vector and the atmospheric particles velocity lie in the XOZ plane: \( \hat{V} = \{ V_x, 0, V_z \}, \quad k = \{ k_x, 0, k_z \} \). Then the conductivity tensor is [25]:

\[
\hat{\sigma} = \begin{pmatrix}
\sigma_p & \sigma_h & 0 \\
-\sigma_h & \sigma_p & 0 \\
0 & 0 & \sigma_t
\end{pmatrix}, \quad (11)
\]

where \( \sigma_p \) is Pedersen’s, \( \sigma_h \) is Hall’s, and \( \sigma_t \) is the parallel component of the conductivity tensor. The altitudinal profile of the components is shown in Figure 2. According to (10), the dynamo-current \( j_{dyn} \) consists of a Hall component parallel to the air particles velocity \( j_{dyn}^{h} = -\sigma_h \hat{V}_x B \), and a Pedersen component perpendicular to the particle velocity. Because of \( \nabla \cdot j_{dyn}^{h} = \nabla \cdot j_{dyn}^{p} \), the Pedersen current does not cause charge separation, and on this basis is not further considered.

As can be seen from Figure 2, the Hall conductivity has a sharp maximum in the altitude range of \( z_d = 90 - 130 \) km (dynamo-layer), where mainly the dynamo-current jet is concentrated. Due to the small thickness of the dynamo-layer in comparison with GW vertical wavelength, the estimated dynamo-current integral value is:

\[
I_{dyn} = \int_0^\infty j_{dyn}^{h} dz = \hat{V}_d B \int_0^\infty \sigma_h dz = \Sigma_h \hat{V}_d B, \quad (12)
\]

where \( \Sigma_h \) is integral Hall conductivity of the atmosphere, and \( \hat{V}_d \) is neutral particles velocity in the dynamo-layer.

The current source (12) is partially shunted by a conductivity current with density \( j_{cs} = \sigma_p E_x \), and is partially closed by means of a parallel current \( j_z \), which is going along open magnetic lines to infinity, and at closed lines to a magnetically conjugate ionosphere. The equivalent electrical circuit generated in this way is shown in Figure 9. The calculation of the circuit parameters is carried out in a standard way adopted in the ionosphere electrodynamics [25].
The integral density of transverse current (dynamo current minus conductivity current, see Figure 9) is:

$$ I_x = I_x^{\text{dyn}} - I_x^c = \frac{\Sigma_x \Sigma_h}{\Sigma_p + \Sigma'_p} \tilde{V}_x B. \tag{13} $$

The transverse electric field due to a voltage drop when current flows through the space circuit:

$$ E_x = \frac{\Sigma_x}{\Sigma_p + \Sigma'_p} \tilde{V}_x B. \tag{14} $$

The magnetic field variation generated in the dynamo-layer (for the model of a flat current sheet):

$$ \tilde{B}_y = 0.5 \mu_0 I_x. \tag{15} $$

The density of the parallel current injected into the magnetosphere from the dynamo-layer:

$$ j_z = -\frac{\partial}{\partial x} I_x \Rightarrow |j_z| = k_x I_x. \tag{16} $$

In these formulas $k_x$ is the GW horizontal wavenumber, and $\Sigma'_p$ and $\Sigma_p$ are the integral Pedersen conductivities in the zone of the dynamo-current generation and in the magnetically conjugate ionosphere (the sum $\Sigma_p + \Sigma'_p$ should be interpreted as the integral conductivity of entire current-carrying tube of the Earth’s magnetic field).

Thus, the value of the ionosphere response to the incoming GWs depends both on the integral Hall conductivity $\Sigma_h$, and on the conditions of the electrical circuit closing in the conjugate ionosphere. In the case of $\Sigma'_p = 0$ (idle mode), the magnetic variation is absent, and the electrical variation is maximum. Opposite behavior corresponds to the case $\Sigma'_p \rightarrow \infty$ (short circuit). With realistic parameters of the atmosphere and GW: $\Sigma_h \sim \Sigma_p \sim \Sigma'_p = 20$ Cm (day), $B = 5 \times 10^{-4}$ T (mean latitudes), $\tilde{V}_{sd} = 20$ m/s (moderate amplitude GW), $\tilde{V}_x B = 45 \times 10^{-6}$ Torr (mean latitudes), $B = 5 \times 10^{-4}$ T (mean latitudes), $\tilde{V}_{sd} = 20$ m/s (moderate amplitude GW).
\[ k_z = \frac{2\pi}{\left(6 \times 10^4\right)} \text{ m}^{-1} \]

we find \( E_z = 1 \text{ mV/m}, \quad B_y = 6.5 \text{ nT}, \quad j_z = 1 \mu\text{A/m}^2 \)

Equations (13)-(16) have to be taken as estimates. They are compiled for the case of the vertical direction of the Earth’s magnetic field (when the conductivity tensor has the form (11)). Analysis of specific experimental situations requires mathematical modeling, taking into account the angle of inclination of the magnetic field, the angle of the direction of propagation of the GW with respect to it, and the distribution of the Hall and Pedersen conductivities in the current-carrying magnetic field tube.

The sketch of atmosphere-ionosphere-magnetosphere perturbation created by GW is given in Figure 10. At heights from the Earth’s surface to the dynamo-layer, we are dealing with a wave of neutral atmosphere (in the figure it is depicted by a wave packet at the bottom right). In the dynamo layer, the wave process is splitting. GW continues to propagate to the left upwards (beyond the limits of the figure) and its dissipation increases with altitude. The final attenuation of the atmospheric branch of the wave process occurs at altitudes 200 - 350 km (depending on the level of solar activity and the spectral composition of GW). Besides, the part of GW energy is spent at generating electromagnetic disturbance, whose horizontal structure reproduces GW profile in the dynamo-layer. Through a parallel current, this structure is transmitted upward into the magnetosphere, experiencing a geometric transformation associated with the divergence of the Earth’s magnetic field lines.

In the F-region of the ionosphere (above 150 km), where the movement of charged particles is controlled by a magnetic field, and not by collisions with neutral particles, the electric field (14) produces a transverse plasma drift with a velocity:

![Figure 10. Sketch of the Geospace disturbance generated by GW propagation.](image-url)
We emphasize that relation (17) characterizes a new mechanism for generating a traveling ionosphere disturbance (TID) not described earlier in the literature. The existing theory interprets the TID as a result of direct involvement of ions in the oscillations of neutral particles [5] [45] [46] [47] [48]. It follows that there can be no TID where there is no GW—away from GW wave packet or higher than the level of GW dissipation. But such a conclusion contradicts to extensive set of satellite data indicating that TID phenomena cover the whole range of ionospheric heights, including in the exosphere (above 300 - 500 km), where all collective movements of the neutral atmosphere cease due to the absence of neutral particles collisions with each other. Instead we show that crossing the dynamo-layer GW generates forced oscillations of the entire Earth’s magnetic field tube. Moreover, relation (17) correctly explains both the TID amplitudes—the velocities of charged and neutral particles are comparable to each other—as well as the non-local nature of GW and TID coupling.

So, several types of electromotive forces act on the ionosphere: atmospheric electricity, discussed in the previous section; dynamo-effect due to the neutral atmosphere motion at the heights of E-region, discussed in this section; and penetration from above of the electric fields from the magnetosphere [26]. Comparison of the efficiencies of these factors is given in Table 2.

### 4. Observational Data

The penetration of GWs from the lower layers of the atmosphere to ionospheric heights is an established experimental fact, confirmed by numerous observations when surface sources of GWs were certainly known (nuclear explosions, catastrophic EQs, etc.) [6] [8]. As an additional example, we present the results of long-term ground-based meteorological and magnetometric observations carried out at the Ukrainian Antarctic station “Vernadsky” (Galindez Island) and at a magnetically conjugate point in the United States (near Boston) [44]. It has been established that the waveforms of GWs, which were recorded as variations of meteorological parameters associated with the passage of weather fronts, with a certain delay are reproduced in the Earth’s magnetic field variations and then transmitted as magnetic disturbances to the magneto-conjugate ionosphere. This

<table>
<thead>
<tr>
<th>Table 2. Quasi-stationary electromagnetic disturbances generated in the E-region by space and surface sources.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Source</strong></td>
</tr>
<tr>
<td>Field of magnetospheric convection (polar caps)</td>
</tr>
<tr>
<td>Atmospheric tide (low latitudes)</td>
</tr>
<tr>
<td>GW (daytime hemisphere)</td>
</tr>
<tr>
<td>Fields of the Earth and weather systems</td>
</tr>
</tbody>
</table>
is a clear confirmation of the GW penetration to the heights of at least the dynamo-layer. The presented theory of ionosphere-magnetosphere current systems generation was developed namely based on these observations.

Other convincing example of seismo-ionosphere coupling efficiency through GW is the registration of tsunami waves movement in variations of the total electron content of the ionosphere (TEC) [49] [50]. Tsunami in the open ocean is a smooth and low water surface lifting with the speed about a fraction of centimeters per second—it is a very unproductive generator of atmosphere oscillations (in comparison, for example, with surface thermal anomalies). At the same time, the TEC value depends mainly on the electron concentration at altitude of ~250 km, where only a small part of GW spectral power penetrates (see the theoretical part of our work). Thus, even such a weak GW source as tsunami, acting on such an inappropriate parameter as TEC, gives rise to a clearly detectable ionospheric effect.

Analysis of DE-2 satellite data. During the 1970s and early 1980s, a series of low Earth orbit satellite missions was implemented to record in-situ the parameters of the upper neutral atmosphere. The largest data set was obtained from DE-2 satellite operating at altitudes of 250 - 500 km in the period 1981-1983 in the conditions of high solar activity.

DE-2 satellite carried a set of instruments including Neutral Atmosphere Composition Spectrometer (NACS) and Wind and Temperature Spectrometer (WATS) which provided the measurement of a full set of hydrodynamic quantities of the neutral atmosphere—concentration, temperature and velocity of the gas components. This opened the possibility of calculating the spectral composition of atmospheric perturbations. Specifically, from the atmospheric particles vertical velocity \( \dot{V}_z \) (directly measured by WATS) and vertical displacement of particles \( \delta z \) (calculated from density variations measured by NACS), one can estimate the wave frequency \( \omega \approx \dot{V}_z / \delta z \) [51]. Further, wave vector components can be estimated using theoretical relations given in [52]. The analytical basis for calculations is the GW theory [3] [53] [54]. Algorithms for satellite data processing in order to decompose them into trend, wave process, and noise are described in [55] [56] [57].

Figure 11 shows the GW waveform registered during one of DE-2 orbits referenced to the globe. This waveform is typical; many similar diagrams were obtained for different values of longitude and local time. It can be concluded that the planetary distribution of GW consists of active high-latitude regions (northern and southern), where the thermosphere is strongly perturbed, and a quiet low-latitude region. The conditional boundary of the regions is at geomagnetic latitudes ~40 - 50 degrees [58] [59] [60].

According to these data, auroras are constantly acting sources of GW generation. From auroral ovals, GWs propagate upward in latitude, filling the polar caps, and towards the equator, gradually attenuating at mid-latitudes. The disturbance in the active regions is so strong that makes it impossible to register GWs from surface sources. In a quiet region, on the contrary, we can extract
Figure 11. GW waveforms in oscillations of O and N₂ neutral gas densities. Dark gray background indicates auroral oval, white and light gray—day and night. The geomagnetic field inclination contours and the South magnetic pole are shown.

them [60]. Since the Mercator projection noticeably distorts the geographical proportions on the map, let us specify that the quiet thermosphere occupies ~70%, and both active regions ~30% of the planet’s surface. And majority of EQ-prone areas are just in quiet zone, which confirms the expedience of the research.

Further, the original results of DE-2 data processing are presented. Figure 12 shows the location of the GW intensity bursts (wave packets) detected in the quiet region during the period November 1982-February 1983, and the location of the epicenters of EQs that occurred during the same period are shown. One can see that GW localization regions are mainly close to seismic ones. Taking into account that in the process of propagation GWs are moving for thousands of kilometers from the source, we can expect that they are caused namely by EQ. To test this assumption, we applied to the data analysis the epoch superposition method—a powerful statistical tool that allows us to reveal hidden relationships even against the background of strong interference. The set of events—EQ and GW—was selected according to the following criteria. Only strong EQs with magnitudes M > 4.5 were taken into account (which is the assumption common in such studies). GW waveforms were the subject of theoretical analysis with the
aim of calculating their spectral characteristics (frequencies, wavelengths, propagation velocities). Further, waves with phase velocity exceeding the speed of sound in the lower atmosphere (~300 m/s) were discarded; such GWs cannot be physically generated by surface sources. In addition, were discarded the couples EQ–GW in which GWs propagated towards the epicenter (according to information on the direction of wave propagation along the satellite’s orbit, determined from the order of the alternation of O and N₂ gases oscillations [61]).

Figure 13 shows a cause-and-effect diagram, in which all EQs are placed at the origin, and GWs are located in the time-distance coordinate system. The cluster of GW is distinguished at an average distance of 6000 km from the epicenter and at times +/- 5 hours relative to the moment of the EQ. We deliberately expanded the time window to two days to demonstrate the rarity of GW outside the cluster. The region of positive times in the diagram we traditionally call the region of responses, although “after” does not necessarily mean “due to”. In this example, the part of the cluster related to positive times lie inside the supersonic cone, and, in our opinion, cannot be associated with an EQ at time zero.

Figure 12. Georeferencing of GW bursts (solid segments) along the orbit of DE-2 and EQ epicenters (circles) for the period November 1982-February 1983. The inlet shows GW waveform.
We believe that all GWs in the cluster—both at negative and positive times—are generated by processes preceding the EQ, although some of the waves reached the satellite after the EQ. To the right of the sound cone there are several GWs with propagation speed of 100 - 150 m/s. We treat them as responses to the EQ shock.

*DEMETER satellite data.* A special goal of the DEMETER mission (2005-2010) was to study the ionospheric EQ precursors. The height of the DEMETER orbit was 650 km, which significantly exceeds the maximal height of GW dissipation. Therefore, TIDs recorded by this satellite and caused by influences from below, should be interpreted as result of magneto-hydrodynamic perturbation during the interaction of GW with dynamo-layer of the ionosphere (see the previous section).

The statistical relationship between variations of ion density and EQ was investigated in [41] based on the data of Ion Analyzer Probe (IAP) aboard this satellite. The set of events was selected according to following conditions: EQ magnitude M > 4.8, satellite distance from the epicenter < 1500, Kp < 3. A time window was not specified, but the authors note that on more than a two week interval, any relationships of EQs and ionosphere disturbances are lost.

*Figure 14* shows an example of the localization of ionospheric disturbances selected in this way relative to the epicenter of a future EQ. Since the horizon of cause-effect analysis was limited to 1500 km, we cannot correlate these data with the diagram in *Figure 13*. We note, however, that the circular structure of the perturbed zone is a characteristic feature of GW.

*Figure 15* (courtesy of Michel Parrot, borrowed from personal presentation) shows a statistical diagram of the EMR generation in the epicentral regions of EQ. The whole data archive of IMSC experiment (Instrument Magnetometre Search Coil) onboard DEMETER satellite (about 9000 hours of measurement sessions for 15,500 orbits during 6.5 years) has been processed using the epoch superposition method. The selected EQs set was limited by conditions: magnitude
Figure 14. Localization of ion density disturbances (circles) relative to the epicenter of a powerful EQ (star) according to DEMETER data (adapted from [41]).

Figure 15. Change in the plasma emissions intensity (magnetic component) relative to the EQ moment (M. Parrot, private communication). Symbol $\sigma$ on the gray scale means statistical dispersion.

M > 5, hypocenter depth < 40 km. Let us pay attention that in addition of the supposed EQs precursors zone (~2 - 5 hours before EQ) also the area ~10 hours after the mainshock is filled with a broadband noise. The only mechanism of postseismic-ionspheric interaction which may be characterized by such a delay time is the GW propagation.

5. Conclusions

The theoretical and experimental patterns of the GW impact on the ionosphere from surface sources (seismogenic, weather, etc.) are generally consistent with
each other. But in detail, both pictures require further study and matching. It should be stated that, even having all the information about the impending EQ (taken at least retrospectively), we are not able to calculate satisfactorily the characteristics of its ionospheric precursor. Still more we are far from solving the inverse problem—by the observed parameters of an ionospheric disturbance to predict the parameters of a future EQ.

In our opinion, the further efforts of theorists should focus on the analysis and mathematical modeling of well-diagnosed seismic-ionospheric events (“case study” of “reference events”). As for future experiments, they should be planned in order to verify the theory—its unequivocal confirmation or refutation. Otherwise, the scope will open again for the creation of hypotheses that in the field of ionospheric EQ precursors have been multiplying for several decades.

Summarizing, we can conclude that the idea of GW as an agent for transferring disturbances from surface sources to the ionosphere makes it possible to explain:

− Transportation of significant energy fluxes modifying the median parameters of D- and E-ionosphere regions (unlike to the electromagnetic radiation and quasi-stationary electric currents);
− Independence of the ionospheric response range from the intensity of the source as a result of nonlinear limitation of the GW growth with height;
− Shift of ionospheric disturbance occurring thousands of km away from the source. The time lag of about half-hour to more than ten hours;
− Transformation of the wave process in neutral atmosphere into forced magneto hydrodynamic oscillations of the ionosphere and magnetosphere.

Acknowledgements

There is authors’ duty and privilege to thank Dr Vira Pronenko and Dr Fedir Dudkin for fruitful discussion and valuable comments which allowed us to improve considerably the paper.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References


An Evidence on the Lithosphere-Ionosphere Coupling in Terms of Atmospheric Gravity Waves on the Basis of a Combined Analysis of Surface Pressure, Ionospheric Perturbations and Ground-Based ULF Variations. *Journal of Atmospheric Electricity*, **33**, 53-68. [https://doi.org/10.1541/jae.33.53](https://doi.org/10.1541/jae.33.53)


