IONOSONDES
and the measurements they make

Bodo Reinisch
Lowell Digisonde International & University of Massachusetts Lowell
Email : Bodo.Reinisch@digisonde.com

International Reference Ionosphere 2019 Workshop
COSPAR Capacity Building Workshop
Frederick University, Nicosia, Cyprus, 2-13 September 2019
Background

- Correct use of measured data is impossible without a full understanding of how the measurements are made. This is especially true for ionosonde measurements.
- Ionosondes have been the workhorse of ionospheric research since the first HF pulsed soundings were made by Breit and Tuve [1925].
- Ionosonde measurements (ionograms) serve as reference (truth) for other ionospheric measurements like:
  - Incoherent Scatter Radar
  - Space-borne in situ
  - GNSS Radio Occultation
- Why are ionogram measurements considered “the truth”?  
  - Vertical HF sounding is a simple technique that directly measures the plasma frequencies in the ionosphere requiring no calibration and no assumptions.
- But why are ionograms so difficult to interpret?  
  - They are not, when recorded with modern ionosondes
  - Automatic scaling and real time analysis provide electron density profiles in real time
An array of receive-antennas is required to measure the arrival angles of the echo signals.
An array of 4 receive antennas forms an interferometer to distinguish between vertical and off-vertical echoes. Crossed-loop antennas identify O- and X-wave echoes.
The ionosonde transmits RF pulses vertically upward, stepping through the frequency range of nominally 0.1 to 30 MHz. As the transmitter frequency increases, the RF wave penetrates deeper into the ionosphere until it reflects at the height where:

\[
\text{radio frequency} = \text{plasma frequency} \quad \text{ (for the O-wave)}
\]

- As the RF pulse approaches the reflection height, its group velocity becomes smaller, approaching zero at the reflection height, then it becomes negative, i.e., pulse propagates downward.
- Eventually, a transmitter frequency is reached at which the radio wave penetrates the layer.
  - For O-waves, penetration occurs when the transmitter frequency exceeds the layer’s peak plasma frequency. This frequency is called the critical or penetration frequency of the layer, e.g., \( \text{foF2} \) for the F2 layer.
  - For X-waves, the penetration frequency is higher than that of the O-wave by \( \approx \frac{1}{2} f_B \) (\( f_B = \text{electron gyro frequency} \)).
HF specular reflection vs UHF Scatter

Snell’s law

Signals reflected back to Ionosonde
Signals refracted away from Ionosonde

~ 0.6 deg

Ionosphere

UHF Radar Beam

Incoherent scatter

Ionosonde

ISR
HF Waves in the Ionosphere

• Can electromagnetic energy propagate in an ionized medium imbedded in the geomagnetic field?
  We know from experience that radio waves can propagate in and through ionospheric layers.

• What kind of RF waves can propagate in the ionosphere?
  We look for solutions in form of progressive plane waves. If such waves exist, they must satisfy Maxwell’s equations, as well as the constitutive relations for the medium.
Solving Maxwell’s Equations for homogeneous magnetoplasma

Maxwells’ Equations

Gauss’s Law \( \nabla \cdot \mathbf{D} = \rho = 0 \)
no net charge density: \( \rho = 0 \) in a “neutral” plasma

No magnetic charge \( \nabla \cdot \mathbf{B} = 0 \)

Faraday’s Law \( \nabla \times \mathbf{E} = -\frac{\delta \mathbf{B}}{\delta t} \)

Ampere’s law \( \nabla \times \mathbf{H} = \mathbf{J} + \frac{\delta \mathbf{D}}{\delta t} \)

\[ \mathbf{B} = \mu \mathbf{H} \]
\[ \mathbf{D} = \varepsilon \mathbf{E} \]
Coupled differential equations

Equation of motion:
\[ m \frac{du}{dt} = eE + eu \times Bg + mvu \]

Ampere’s Law becomes:
\[ \nabla \times \mathbf{H} = J + \frac{\delta D}{\delta t} = Neu + \frac{\delta(\varepsilon_0 E)}{\delta t}, \text{ and} \]
\[ \nabla^2 E = \varepsilon_0 \mu_0 \frac{\delta^2}{\delta t^2} E \quad \text{and} \quad \nabla^2 B = \varepsilon_0 \mu_0 \frac{\delta^2}{\delta t^2} B \]

Assume progressive plane wave solutions exist:
\[ E(r, t) = E_0 \exp \{i(\omega t - k \cdot r + \varphi_0)\}, \text{ then:} \]
\[ \nabla \times \mathbf{H} = i\omega (Ner + \varepsilon_0 E) \]
since \[ \frac{\delta r}{\delta t} = \mathbf{u} \text{ and } \frac{\delta}{\delta t} \equiv i\omega \text{ for a plane progressive wave} \]

\( \nu \) is the collision frequency of electrons with neutrals causing absorption

For an “electron gas”:
\[ e = -1.6 \cdot 10^{-19} \text{Coulomb} \]
Force on free electrons

The force exerted on the free electrons by the electric field of the HF radiowave displaces the electron from its rest place. For HF radiowaves, the displacement of the ions can be neglected (to first order); unlike the electrons the ions are heavy and barely move. The magnetic Lorentz force $e\mathbf{u} \times \mathbf{B}$ is also negligible compared to the electric force $e\mathbf{E}$. 
Solutions for the wave equations

• The path of a radiowave is affected by the medium through which it travels. The *refractive index* \( n = \frac{v_{ph}}{c} \) for HF radiowaves in the ionosphere is governed by the *electron concentration* \( N \) and the *magnetic field* \( B_g \) of the medium.
• The index \( n \) is a function of the radio frequency, wave polarization, and wave direction (*anisotropy*) of the EM wave.
• Only RF waves with *specific wave polarizations* can propagate, usually referred to as the *characteristic waves*.
  
  – Usually two characteristic waves exist, called *ordinary* and *extraordinary* (or O and X) modes.
  
  – A third characteristic wave can exist, the so-called *Z-mode*, but it is rarely observed in the bottomside ionosphere (usually at high latitudes). [see Benson et al., 2006.]
Characteristic Wave Polarizations

Using the set of equations shown before, one can show (after some exhausting arithmetic) that plane waves $E(r, t) = E_0 \exp[i(\omega t - k \cdot r + \varphi_0)]$ can propagate in a magnetoplasma and maintain its polarization, if they have one of two characteristic wave polarizations.

- The characteristic waves have right- and left-hand elliptical wave polarization, which at most latitudes are almost circular.
- Except at the magnetic equator where the two characteristic polarizations are linear: parallel to $B_g$ for the O-wave, and perpendicular for the X-wave.
The indices of refraction $n_{\pm}$ are given by the **Appleton-Lassen** formula:

$$n_{\pm}^2 = 1 - \frac{X}{1 - iZ} - \frac{Y_T^2}{2(1 - X - iZ)} \pm \sqrt{\frac{Y_T^4}{4(1 - X - iZ)^2} + Y_L^2}$$

where

$$X = f_N^2 / f^2; \quad Y = f_B / f; \quad Z = \nu / 2\pi f$$

and

$$f_N^2 = \frac{e^2}{4\pi^2 \varepsilon_0 m} N \approx 80.6 N \quad (\text{with } N \text{ in } m^{-3}, \quad f_N \text{ in } Hz)$$

$$f_B = \frac{eB_g}{2\pi m} \approx 1.4 \text{ MHz}$$


*often called “Appleton-Hartree” formula, but Hartree (1931) had included an incorrect dielectric polarization term. Lassen had derived the correct $n_{\pm}$ in 1927.*
Absorption of Radiowaves

\[ n_\pm^2 = 1 - \frac{X}{1 - iZ - \frac{Y_T^2}{2(1 - X - iZ)} \pm \sqrt{\frac{Y_T^4}{4(1 - X - iZ)^2} + Y_L^2}} = (\mu_\pm - i\chi_\pm)^2 \]

For a wave propagating in the \( x_1 \) direction:

\[ E = E_0 \exp \{\omega t - k x_1 + \phi_0\} \]

\[ k = \frac{\omega}{v_{ph}} = \frac{\omega_n}{c} = \frac{\omega}{c} (\mu - i\chi) = k_0 (\mu - i\chi) \text{ since the phase velocity } v_{ph} \text{ is } v_{ph} = \frac{c}{n}. \]

Define the "absorption coefficient": \( \kappa = \frac{\omega}{c} \chi = k_0 \chi, \quad k_0 = \frac{2\pi}{\lambda_0} \)

\[ E = E_0 e^{-\kappa x_1} e^{i(\omega t - \mu k_0 x_1)} \]

\[ \kappa_\pm \approx \frac{e^2}{2\varepsilon_0 mc} \frac{1}{\mu} \frac{Nv}{(\omega_\pm \omega_g)^2 + v^2} \]

(e.g., Davies, 1990, Ionospheric Radio, Peregrinus Press)

Note: The X-wave (minus sign) has larger absorption since \((\omega - \omega_g) < (\omega + \omega_g)\).
"Group" Velocity and Refractive Index

In a dispersive medium, the phase and group velocities differ. The phase velocity is (neglecting collisions)
\[ v_{ph} = \frac{c}{\mu}. \]
Since \( \mu < 1 \) in a plasma, \( v_{ph} > c \)

The group velocity in a dispersive medium is
\[ v_g = \frac{d\omega}{dk}. \]

The group refractive index is
\[ \mu' = \frac{c}{v_g} = c \frac{d\kappa}{d\omega} = \frac{d}{d\omega} (\mu\omega) = \mu + \omega \frac{d\mu}{d\omega}. \]

Neglecting the magnetic field and collisions \((Y=0 \text{ and } Z=0)\): (just to simplify the discussion here)
\[ \mu = \sqrt{1 - (f_N/f)^2}. \]
Then
\[ \mu' = \frac{d}{df} (\mu f) = \frac{1}{\mu} \quad \text{or} \quad \mu\mu' = 1 \quad \text{Therefore} \quad v_g v_{ph} = c^2 \]

Note that \( v_g \to 0 \) at the reflection point where \( \mu = 0 \).

Important:
Directions of \( v_{ph} \) and \( v_g \) are different when using the plane wave description!
Huang and Reinisch [2012] showed that using spherical wave solutions makes \( v_{ph} \parallel v_g \).

Digital ionosondes became available in the 1970s, starting with the Lowell Digisonde-128 in **1969**. The development of the Digisonde® family [Reinisch et al., 2009] is portrayed below. Several other digital ionosonde types became available in the last 35 years (KEL-72/Aus, Dynasonde/USA, CADI/Canada, VIPIR/USA, Parus/Russia, AIS/Italy, VISRC/Poland, SPICE/Aus, CAS-DIS/China, ...).
Ionograms

- The basic product of an ionosonde is the *ionogram*, which plots the received echo amplitude in a *virtual height vs frequency* frame
  - The virtual height is \( h'(f) = \frac{1}{2} \cdot \tau(f) \cdot c \), where \( \tau \) is the measured pulse travel time, and \( c \) is the free-space speed of light
- Modern digital ionosondes also
  1. **Measure:**
     - Wave polarization
     - Angle of arrival
     - Doppler frequency
  2. **Automatically “scale”** the ionogram in real time
  3. **Calculate the vertical electron density profile** in real time
  4. **Ingest ionograms and scaled data in global data centers** in real time
Automatic Scaling of Ionograms

Ionosondes – Bodo Reinisch

Autoscaled:

- \( f_{\text{F}2} = 5.49 \)
- \( f_{\text{F}1} = 4.86 \)
- \( f_{\text{F}1p} = 4.56 \)
- \( f_{\text{O}} = 2.96 \)
- \( f_{\text{O}p} = 3.23 \)
- \( f_{\text{I}} = 6.28 \)
- \( f_{\text{O}e} = 2.95 \)
- \( f_{\text{min}} = 1.65 \)

- MUF = 18.79
- M = 3.424
- I = 3000

- \( h^*_{\text{F}2} = 185 \)
- \( h^*_{\text{F}1} = 235 \)
- \( h^*_{\text{E}} = 97 \)
- \( h^*_{\text{Es}} = 98 \)

- \( z_{\text{F}2} = 225 \)
- \( z_{\text{F}1} = 195 \)
- \( z_{\text{E}} = 106 \)
- \( y_{\text{F}2} = 65 \)
- \( y_{\text{F}1} = 48 \)
- \( y_{\text{E}} = 15 \)
- \( D_0 = 60.1 \)
- \( D_1 = 2.39 \)

- C-level = N/A

D  100  200  400  600  800  1000  1500  3000  [km]
MUF  6.2  6.3  6.6  7.0  7.7  8.7  11.5  18.8  [MHz]
Spread-F Conditions in Auroral Region

Autoscaled data

Oblique echoes:
- from East
- from South-South-East

X-data

Vertical O-trace

N(h)-profile

Conventional ionogram

Same ionogram without pixel identification

Ionosondes – Bodo Reinisch  
IRI 2019 Workshop, Nicosia Cyprus
Equatorial Plasma Bubbles over Cachimbo
Contour plots of equal plasma frequency $f_N$ derived from vertical-incidence ionograms, measured every 5-min. Color code gives plasma frequency in MHz (repeating every 1MHz). [Courtesy D. Altadill, Ebro Obs]
The ionosonde measures the virtual heights $h'(f)$ of the reflection.

**Question:**
How is the **real** height (or **true** height) of the reflections calculated? Knowing the real reflection height $h(f)$ for every frequency $f$ of the O-wave defines the plasma frequency vertical profile since the O-wave reflects at the height where

$$f(Hz) = f_N \approx 9\sqrt{N}[m^{-3}]$$

**Answer:**
A number of extensively tested “true height inversion” algorithms are available.
Plasma frequency profiles $f_N(h)$

Profile $f_N(h)$ (real height)
If the measured propagation time of the pulse from the ground to the reflection point in the ionosphere is $T(f)$, we can define a virtual reflection height $h'(f)$:

$$h'(f) = \frac{1}{2} c T(f),$$

which assumes, incorrectly, that the pulse velocity is the speed of light $c$. In reality, the pulse travels with the group velocity $v_g(f) = \frac{c}{\mu'}$.

The total travel time is

$$T(f, h_r) = 2 \int_0^{h_r(f)} \frac{dh}{v_g(f, h)} = \frac{2}{c} \int_0^{h_r(f)} \mu' dh.$$

Therefore:

$$h'(f) = \int_0^{h_r(f)} \mu'(f, h) dh \quad \text{where} \quad \mu' = \frac{d}{d\omega} (\mu \omega).$$

Here $\mu(f, h)$ is the index of refraction (see Appleton-Lassen formula) which changes from $\mu = 1$ (below the ionosphere) to $\mu = 0$ (at reflection). We need to find $h_r$ for each sounding frequency $f$. 
Calculating the Plasma Frequency Profile

The **POLAN inversion** program is often used to calculate the EDP, or $f_N(h)$, from the echo traces in the ionogram [Titheridge, 1988], but there are many other programs.

Today the most-frequently used **real height inversion program is NHPC**, which is applied in all Digisondes, and is also used to process any other digitized ionograms. In NHPC [Reinisch and Huang, 1983], the $f_N(h)$ profile for each layer is represented as:

$$ h = h_m + \sqrt{g} \sum_{i=1}^{5} A_i T_i^*(g) \quad \text{with} \quad g = \ln \frac{f_N}{f_m} \frac{f_s}{f_m} $$

$$ h_{\text{start}} = h_m + \sum_{i=1}^{5} A_i $$

$T_i^*$ = $i$-th order Chebyshev polinominal

$(h, f_N)$ = (height, plasma frequency)

$(h_m, f_m)$ = (height, plasma frequency at layer peak)

$(h_s, f_s)$ = (height, plasma frequency at the layer bottom)

The $A_i$ coefficients are varied until the **recalculated** $h'(f)$ trace matches the measured $h'(f)$ trace.
Validation of ionogram-based N(h) profiles
Several thousand inverted profiles from Digisonde ionograms at Millstone Hill (42.6N,284.5E) and Ramy, Puerto Rico (18.5N,292.9E) were compared with the profiles derived from co-located incoherent scatter radar measurements for half a solar cycle from 1990-1996 [Chen, et al., 1993].

Results
1. The differences in $h(f_N)$ are in the average less than 5 km for F layers
2. The E-F valley is modeled

Uncertainties of $h'(f)$ measurements
~ 1 km in ionosondes with phase-range-measurement technique
[Reinisch et al., 2008]
Assimilation of Ionogram Measurements in the IRI Profile Model

Ionograms measure:
- $f_{oF2}$, $h_mF2$ and $f_{oF1}$, $h_mF1$

IRI F2-layer bottomside profile

$$N(h) = NmF2 \frac{\exp(-x^{B1})}{\cosh(x)}$$

$$x = \frac{h_mF2 - h}{B_0}$$

How to get $B0, B1$ from the ionogram?

Match the IRI F2 profile to the measured F2 profile.

How to get $B0, B1$ from the ionogram?
IRTAM assimilates ionosonde data in IRI
24-hour real-time GIRO ionogram movies

http://giro.uml.edu/IonogramMovies/

Boulder  C.Paulista  Dourbes  El Arenosillo  Ebro  Eglin AFB

Gakona  Hermanus  Irkutsk  Jicamarca  Jeju Is.  Kwajalein Is.

Millstone Hill  Moscow  Norlisk  Pt. Arguello  Pruhonice  RRL-Chilton
From the raw ionogram data set the virtual heights $h'(f)$ for the vertical echoes must be specified, i.e. the ionogram must be “scaled”:

- Find the virtual heights $h'(f)$ for the E, F1, and F2 layer echoes,
- Determine the critical frequencies $f_{oE}$, $f_{oF1}$, and $f_{oF2}$

This used to be a tedious process with the old analog ionosonde recordings. With the advent of digital ionogram recordings came the development of **automatic scaling routines**.

- The most-widely-used one is **ARTIST-5**, which scales the ionograms in the Digisondes, and –using NHPC- calculates the EDP in real time [Galkin et al., 2008; Reinisch and Huang, 1983].
- The correctness of the calculated EDP depends critically on the correct scaling of the ionogram. Of course, automatic scaling makes mistakes (more than a skilled manual scaler makes). Computerized techniques have been developed to efficiently **detect/correct** these scaling mistakes. The GIRO Digisonde community uses SAO-Explorer for this task.
Editing the autoscaling with SAO-Explorer

Grahamstown

Good ARTIST scaling, no editing required

Vertical echoes at Grahamstown

(Oblique ionogram signals from Hermanus)

ARTIST echo trace h'(f)

EDP
Example of wrong ARTIST autoscaling
Topside Sounding

- Ionosondes on polar orbiting satellites at ~900-1400 km altitude can measure foF2 and hmF2, and the topside electron density profiles from hmF2 up to the satellite altitude.
- foF2/hmF2 values from the topside sounders can complement the groundbased GIRO data in regions of scarce GIRO population, especially over the oceans and polar regions.
- Real-time assimilation of the topside sounder data in ionospheric models, like IRI or NeQuick, will significantly increase the accuracy of the model predictions.
ISIS Topside Ionograms
and “TOPIST” autoscaling of ISIS ionograms

ISIS ionograms have no polarization tagging. TOPIST autoscaling successfully scales undisturbed ionograms

Lack of polarization tagging makes it difficult to reliably identify foF2.

[Huang et al., 2002]
Proposed Topside Orbit
Modern Topside Ionospheric Sounder
Double-Probe & Topside–Ionosphere-Sounder, DP-TIS
(Prototype built for AFRL)
Antenna Boom Configuration for DPTIS

- O and X echo identification
- Angle-of-arrival measurements
- Ionogram data records
- Skymap data records
- Relaxation sounding data records
- automatic EDP calculation
- foF2/hmF2
Some References


