The Effect of the Ionosphere on Low-Frequency Subsurface Radio Soundings of Mars

by

Donald A. Gurnett

February 2, 1998

Dept. of Physics and Astronomy
The University of Iowa
Iowa City, IA 52242

Document 97-60002

*A brief report prepared for the Mars Express Sounder team.

In order to provide deep subsurface penetration, the frequency of the Mars Express Sounder must be as low as possible. The ionosphere of Mars imposes two constraints on how low this frequency can be. First, since in an ionized medium electromagnetic radiation cannot propagate at frequencies below the electron plasma frequency, the sounder must operate at a frequency above the maximum electron plasma frequency in the ionosphere. The electron plasma frequency is given by $f_p = 8980 \sqrt{n_e}$ Hz, where n_e is the electron density in cm^{-3}. Second, since the index of refraction depends on frequency, the ionosphere causes a temporal dispersion of the sounder signal. If the sounder frequency is too close to the plasma frequency, the dispersion become large and extremely sensitive to small variations in the ionospheric density. Since small-scale density fluctuations are likely to be present in the ionosphere, the resulting distortion caused by these fluctuations place constraints on how close the sounder frequency can be to the plasma frequency.

To evaluate the first constraint, a plot of the maximum plasma frequency, f_p(Max), is shown in Figure 1 as a function of solar zenith angle. These data are derived from radio occultation measurements of the Martian ionosphere using Mariners 5, 7, and 9, Mars 2 and 3, and Vikings 1 and 2 [from Zhang et al., 1990a, 1990b]. Note that no measurements are available on the dayside at solar zenith angles less than 45°, or on the nightside at solar zenith angles greater than 130°. The solid curves give extrapolations of the upper and lower bounds to the plasma frequency into the region near the subsolar point. These extrapolations are based on the Chapman theory of planetary ionospheres, which predicts that the peak density should vary as $(\cos \theta)^{1/2}$, where θ is the solar zenith angle. As one can see, the maximum plasma frequency varies from about 4.3 MHz near the subsolar point to about 800 kHz on the nightside. Because of range considerations, subsurface soundings must be performed near periapsis. To achieve the lowest possible sounding frequency, a periapsis on the nightside clearly provides the best opportunity to perform subsurface soundings, since the lowest plasma frequencies occur in this region. Unfortunately, we know very little about the structure of the nightside ionosphere at Mars. From our more detailed knowledge of the ionosphere of Venus [Brace et al., 1983], we know that nonmagnetized planets often have deep holes in the ionospheric electron density on the nightside. Indeed some of the nightside radio occultation measurements at Mars show no detectable ionosphere, so it might be possible to perform soundings at frequencies well below 800 kHz in this region.

Next we turn to the constraint imposed by temporal dispersion of the radar signal as it passes through the ionosphere. Dispersion occurs because the index of refraction varies with frequency. In a plasma with no magnetic field, such as at Mars, the index of refraction is given by

$$n = \sqrt{1 - (f_p/f)^2} \quad . \tag{1}$$

As can be seen, when the wave frequency is below the plasma frequency, $f < f_p$, the index of refraction becomes imaginary. A sounder signal incident on the ionosphere in this frequency range is completely reflected when the wave frequency reaches the local plasma frequency. This reflection process is illustrated in Figure 2, which shows a representative profile of the plasma frequency as a function of altitude. As soon as the sounder frequency exceeds the maximum plasma frequency in the ionosphere, f_p(Max), the signal can reach the surface. However, in the process of passing through the ionosphere the signal experiences a phase shift relative due to the presence of the ionospheric plasma. For a signal with a finite bandwidth, the frequency dependence of this phase shift causes different frequency components to be delayed by different times, thereby distorting the phase of the returning signal. For a chirp radar system, such as is being proposed for Mars Express, the deconvolution process fails if the phase shift exceeds roughly 180° over the bandwidth of the chirp. In our proposal we describe a plan to correct for this ionospheric phase shift by constructing a matched filter that is based on a statistical analysis of the strong surface signal. However, this statistical analysis requires averaging over many radar returns, and will likely take 10 to 20 seconds. If the ionosphere changes significantly over this averaging interval, then the matched filter technique fails, since it cannot reconstruct the distorted chirp waveform. To determine the limitations on the system, we need to make a rough evaluation of the phase distortion induced by fluctuations in the ionospheric density.

To evaluate the phase shift caused by propagation through the ionosphere, we start by assuming that the ionosphere consists of a horizontal planar slab with a constant plasma frequency, f_p, and a vertical thickness, L. It is easy to see that the total phase advance that occurs in propagating through this layer is $\phi = kL$, where $k = n\omega/c$ is the wave number, n is the index of refraction, ω is the angular frequency, and c is the speed of light. Since we want to isolate the effect of the plasma, it is convenient to discuss the phase difference $\delta_i\phi = \phi_o - \phi_i$, where ϕ_o is the phase that would exist in free space, and ϕ_i is the phase that exists in the presence of the ionosphere. For reference plots of ϕ_o and ϕ_i as a function of frequency would appear as shown in Figure 3. From the above equations, it is then easy to show that the phase shift due to the presence of the ionosphere is given by

$$\delta_i\phi = \frac{\omega L}{c}[1 - n] \quad , \tag{2}$$

which after substituting Equation 1 for the index of refraction, and $\omega = 2\pi f$, becomes

$$\delta_i\phi = \frac{2\pi f L}{c}\left[\,1 - \sqrt{1 - (f_p/f)^2}\,\right] \quad . \tag{3}$$

Next we must consider the total phase shift due to the ionosphere over the chirp bandwidth, Δf. For this calculation we linearize the above equation about some center frequency, which we will continue to denote by f. It is obvious that the linearization process is not accurate if $\Delta f/f$ is too large, but for our purposes it is probably gives a reasonable good first approximation. Keep only the first term in a Taylor series expansion; it is easy to see that the total phase shift due to the ionosphere across the frequency band Δf is given by

$$\Delta(\delta_i\phi) = \left[\frac{\partial}{\partial\phi}(\delta_i\phi)\right]\Delta f \quad , \tag{4}$$

which works out to be

$$\Delta(\delta_i\phi) = \frac{2\pi L\Delta f}{c}\left[\frac{1}{\sqrt{1 - (f_p/f)^2}} - 1\right] \quad . \tag{5}$$

Note that the total phase shift across the band is simply proportional to Δf, which makes sense, since if the bandwidth is narrow, there is relatively little variation in the index of refraction over the frequency band of the pulse. As Δf increases, the variation in the index of refraction increases and the phase shift increases. It is interesting to note that the multiplicative factor, $2\pi L\Delta f/c$, to the right of the bracket in Equation 5 only depends on the thickness of the ionosphere and the bandwidth. It does not depend on the plasma frequency. The plasma frequency control is completely isolated to the term in the bracket, which is directly linked to the frequency dependence of the index of refraction. For typical parameters, Δf = 1 MHz and L = 50 km the multiplicative factor turns out to be a quite large, $2\pi L\Delta f/c$ = 1,046 radians. This immediately tells us that the phase shift due to the ionosphere is quite large and cannot be ignored. The frequency dependence of the phase shift is illustrated in Figure 4, which shows the phase shift as a function of the ratio of the wave frequency to the plasma frequency. As can be seen, the phase shift remains large up to very high frequencies. A phase shift of less than one radian, $\Delta(\delta_i\phi) \leq 1$, can only be achieved if the wave frequency is more than 23 times the plasma frequency. Such a high sounding frequency cannot be considered for subsurface sounding at Mars. It follows then that the ionosphere always causes a very large phase distortion, which must be removed by some suitable technique.

The presently proposed technique for removing the ionospheric dispersion is to statistically construct a matched filter by suitably averaging a large number of surface returns. Note that the matched filter correction is not a matter of correcting a few degrees of phase shift, such as one might do to correct for a phase shift in the antenna system. At a sounding frequency of twice the peak ionosphere plasma frequency, which is a very realistic value, the total phase shift that must be corrected across the 1 MHz bandwidth of the chirp signal is roughly 170 radians. Whether this can in fact be corrected using the matched filter approach depends entirely on how smooth and constant the ionosphere is during the time that it takes to perform the statistical construction of the matched filter, which I understand to be about 10 to 20 seconds. Of course, if the ionosphere is absolutely smooth and steady, it will work. The problem is that the ionosphere is not likely to be smooth and steady.

To investigate the sensitivity of the phase shift to variations in the ionospheric electron density, it is a simple matter to differentiate Equation 5 with respect to f_p. Computing this derivative, and noting $\delta f_p/f_p = (1/2)\delta n_e/n_e$, where δn_e is the change in the electron density, it is easy to show that the change in the phase is given by

$$\delta\Delta(\delta_i\phi) = \frac{2\pi L\Delta f}{C} \frac{f_p^2/f^2}{(1 - f_p^2/f^2)^{3/2}} \frac{1}{2} \frac{\delta n_e}{n_e} \quad . \tag{6}$$

Note that the phase shift is directly proportional to the fractional change in the electron density. Since one can identify the product $L\delta n_e$ in the numerator, it follows that the phase shift is directly proportional to the perturbation in the electron column density measured vertically through the ionosphere. Assuming that we can only allow a phase shift of about 180° (π radians) during the time it takes to carry out the matched filter analysis, we can then compute an upper limit to the allowed fractional density fluctuation

$$\frac{\delta n_e}{n_e} = \frac{c}{L\Delta f} \frac{[1 - (f_p/f)^2]^{3/2}}{(f_p/f)^2} \quad . \tag{7}$$

This function is shown in Figure 5. As one might expect, the allowed maximum fractional density variation is very small for sounding frequencies near the plasma frequency, but then increases rapidly as f/f_p increases, more or less as one would expect. For a sounding frequency of twice the plasma frequency, $f/f_p = 2$, the maximum allowed fractional density fluctuation is about 1.7

percent. At $f/f_p = 1.25$ the maximum allowable density fluctuation is only 0.2 percent. Whether density fluctuations of this magnitude occurs in the Martian ionosphere over spatial scales corresponding to 10 to 20 seconds of orbital motion (corresponding to 30 km) is unknown. However, based on observations in the ionospheres of Earth and Venus, I would say that fluctuations of this magnitude are quite possible, particularly in regions where gravity waves and auroral disturbances are present.

Frankly, I believe that spatial variations in the ionospheric electron density pose a significant problem for the matched filter technique. There are many things that I do not know about this technique. For example, does the deconvolution process fail gracefully or catastrophically as the phase shift exceeds 180°? If the deconvolution output declines relatively slowly as the phase shift goes above 180°, then maybe it is no problem. To the extent that ionospheric irregularities remain a serious difficulty, I see the following step that can minimize their effect.

(1) Operate in regions with the lowest possible plasma frequency (i.e., on the nightside of Mars).

(2) Reduce the chirp bandwidth. Obviously this reduces the range resolution. But maybe this is a price we have to pay.

(3) Find a way to actively correct for fluctuations in the ionospheric electron density, by using interleaved ionospheric soundings. I am thinking about this, but so far I do not know if any of my ideas would be workable in practice.

References

Brace, L. H., H. A. Taylor, Jr., T. I. Gombosi, A. J. Kliore, W. C. Knudsen, and A. F. Nagy, The ionosphere of Venus: Observations and their interpretation, *Venus*, Univ. of Arizona Press, 779-840, 1983.

Zhang, M. H. G., J. G. Luhmann, A. J. Kliore, and J. Kim, A post-pioneer Venus reassessment of the Martian dayside ionosphere as observed by radio occultation methods, *J. Geophys. Res.*, *95*, 14,829-14,839, 1990a.

Zhang, M. H. G., J. G. Luhmann, and A. J. Kliore, An observational study of the nightside ionospheres of Mars and Venus with radio occultation methods, *J. Geophys. Res.*, *95*, 17,095-17,102, 1990b.

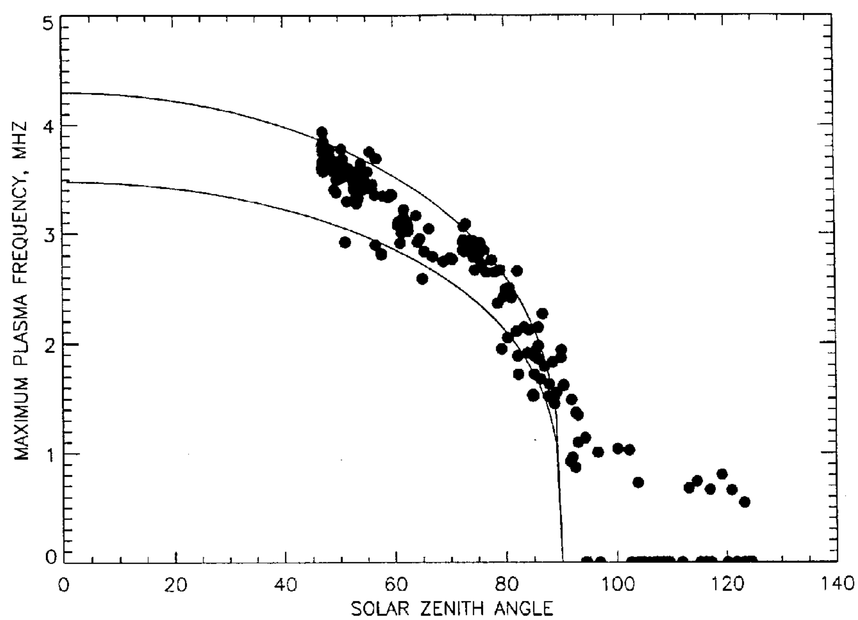

Figure 1

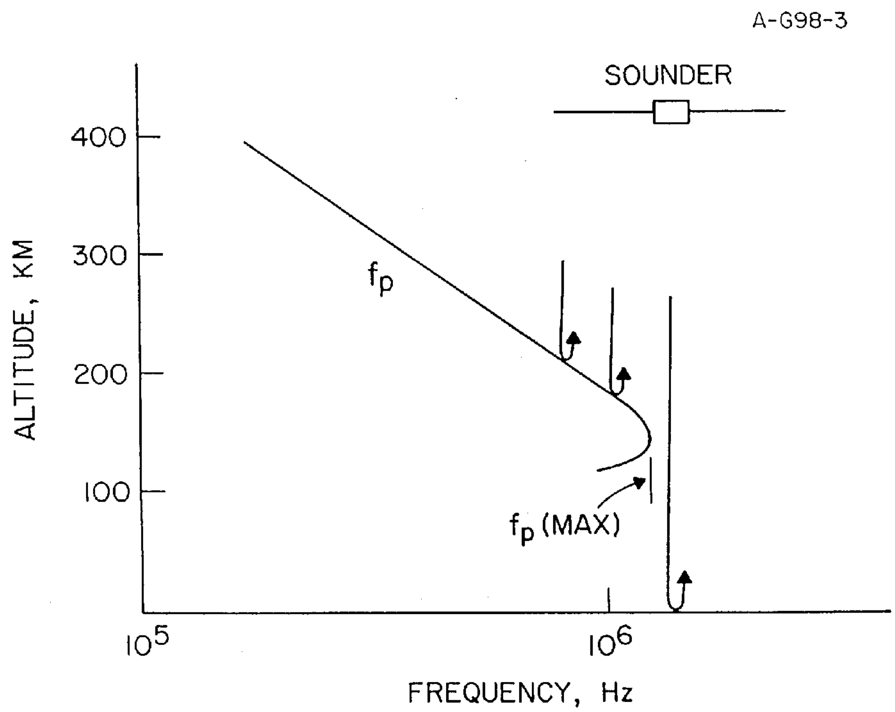

Figure 2

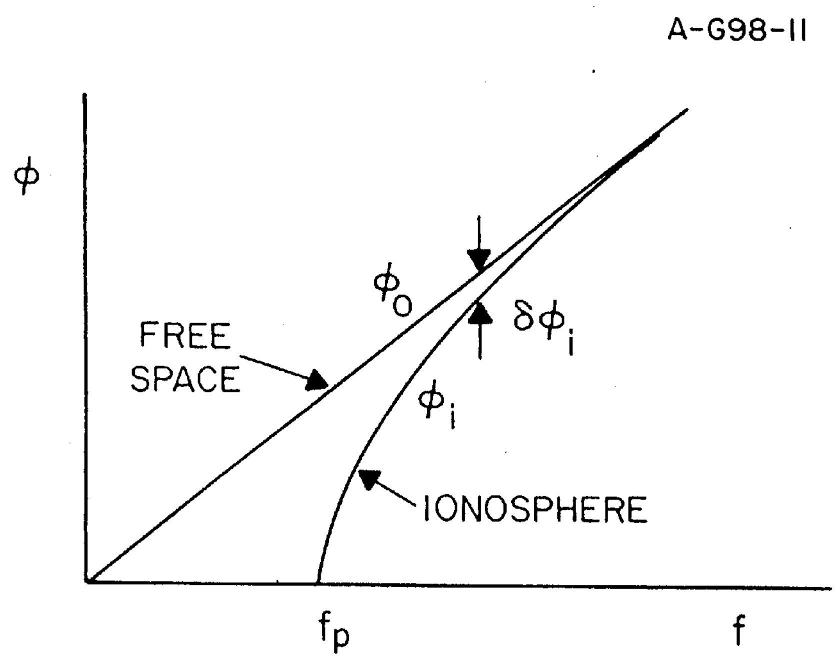

Figure 3

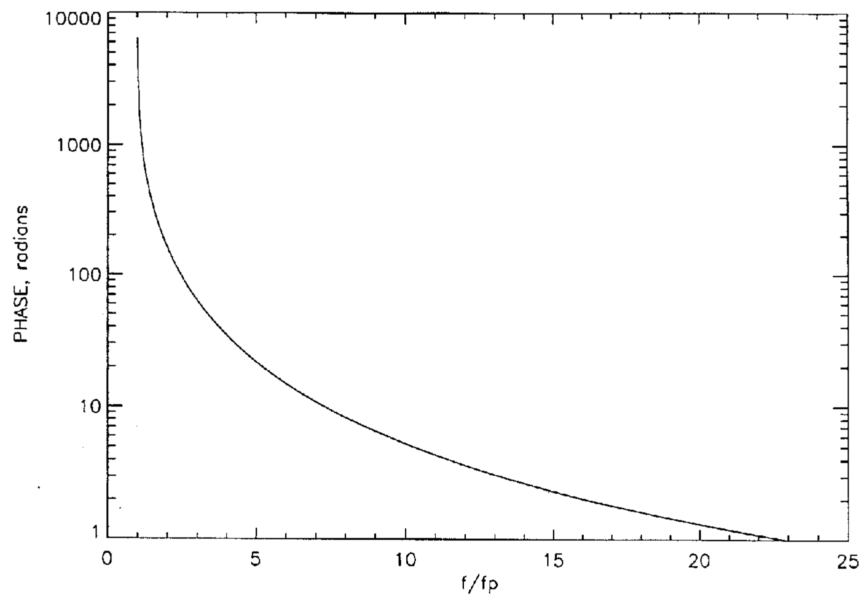

Figure 4

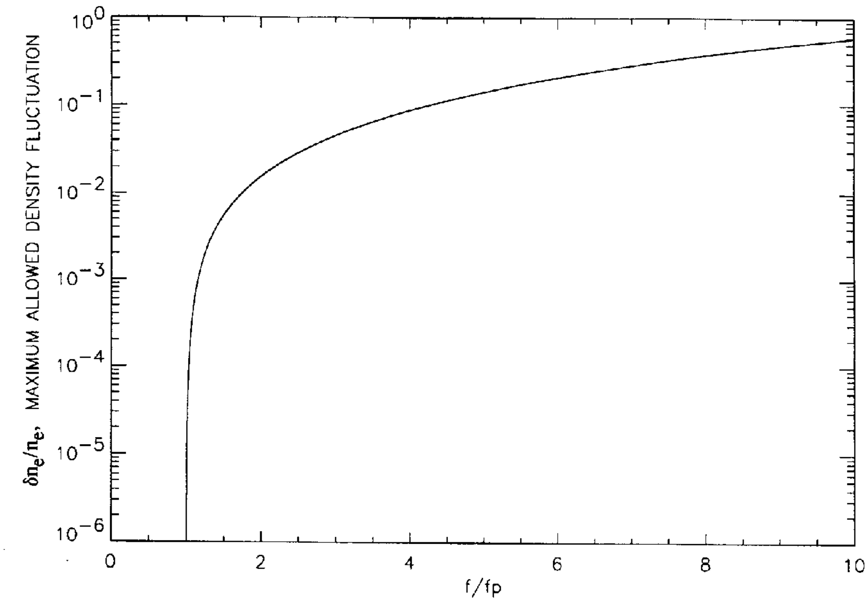

Figure 5