3. Radio Occultation Principles

The radio occultation technique was first developed at the Stanford University Center for Radar Astronomy (SUCRA) for studies of planetary atmospheres. Radio occultation experiments at the Jet Propulsion Laboratory (JPL) have played a prominent role in the NASA program for solar system exploration for more than two decades. They have contributed uniquely to studies of the atmospheres of Venus[8], Mars[9], the gas giants Jupiter, Saturn, Uranus, and Neptune[10], as well as the outer-planet satellites Io, Titan, and Triton.[11] Typically, experiments involved a spacecraft transmitter linked to a terrestrial receiver via a cm-wavelength radio signal. The spacecraft trajectory was selected so that the propagation path from the spacecraft to Earth passed through the planetary atmosphere under study, producing distinctive variations in the amplitude and frequency (or phase) of the received signal.

Fundamentally, the technique relies on the simple fact that a planet's atmosphere acts much like a spherical lens, bending and slowing the propagation of microwave signals passing through it tangent to the surface. The lens effect results from decreasing atmospheric density with altitude. If the positions of transmitting and receiving satellites are precisely known, the "atmosphere delay" can be measured precisely, the time derivative of which (Doppler) can be inverted to give atmospheric density vs. altitude.

For an Earth observing system based on the radio occultation technique, the cost of maintaining a constellation of Earth orbiting satellites transmitting on appropriate frequencies would be dominant. In contrast, the receiving satellites would be relatively inexpensive. It so happens that the GPS exists, is free of charge, and already has 24 satellites transmitting on frequencies suitable for occultation observations. Moreover, by using the GPS to derive the precise satellite positions, overall system complexity (and cost) are further reduced. Thus, there is a strong economic incentive to base an Earth radio occultation observing system on GPS.

3.1 The GPS Navigation System

The GPS is a state of the art satellite navigation system. It consists of 24 operational satellites, 4 in each of six 12 hour, 20,000 km circular orbits, all inclined 55deg.. The resulting constellation produces global coverage 24 hours a day. There is no charge for use of the service and the U S Government has recently issued a policy statement assuring the international community that no user fees will be imposed for at least the next 10 years.

The GPS satellites transmit on two L-band carrier frequencies: 1575.42 MHz (L1) and 1227.6 MHz (L2). Each carrier is phase modulated by a precise ranging code (P code) consisting of pseudo random bit sequences at 10.23 Mb/s. In addition, the L1 carrier is modulated in quadrature with a 1.023 Mb/s pseudo random bit sequence used for the coarse (or clear) acquisition code (C/A code). The transmit time, as kept by the clock onboard each GPS satellite, is precisely known for each bit in the sequence. A GPS receiver identifies the incoming code bits and measures their arrival time, as kept by the receiver clock, with a precision of better than 1% of a bit length (about 1 nsec or 30 cm for the P code). A priori GPS orbital positions and clock offsets between GPS satellites are broadcast to the user along with other information on a 50 bps data message superimposed on the L1 and L2 carriers. The difference between the known transmit time and observed arrival time is a measure of the distance between the satellite and receiver, plus the clock offset between transmitter and receiver clocks, a quantity referred to as "pseudorange." A receiver simultaneously measuring pseudorange to four satellites can instantaneously
determine its three components of position and its time offset from GPS time, typically with an accuracy of 10-15 m and <1 microsecond respectively. Modern receivers can also measure and keep continuous count of carrier phase with a precision of better than .5% of a wavelength (~ 1 mm). Continuous phase can then be used to construct a record of position change with millimeter precision.

For reasons of national security, current U.S. Government policy calls for limiting access to the Precision Positioning Service (PPS), and the accuracy of the Standard Positioning Service (SPS). Two techniques are used to limit the access and the accuracy of GPS: Selective Availability (S/A) and/or Anti Spoofing (A/S). A/S is a process used to deny users access to the full capabilities of the system by encryption of the high rate P code normally required for high precision measurements. When so encrypted, the high rate code is referred to as the "Y code". Unless the user has the required "encryption key" to track the Y code, the user will not have access to the PPS. S/A is believed to involve the deliberate introduction of small, random errors in the satellite ephemeris data broadcast in the almanac, and in the carrier and/or clock frequency transmitted. Uncorrected, S/A can result in position errors on the order of 100 m.

For GPS/MET, access to the highest precision available from GPS is required. However, "Y Code receivers" and encryption keys are not needed. Instead, a "codeless receiver", capable of tracking the L2 carrier phase without explicit knowledge of the Y Code, is used. With respect to S/A, a UNAVCO study has shown that when Double Differencing (described below) is used in conjunction with synchronized receiver clocks, S/A is effectively canceled, just as normal clock and orbit errors are canceled. Therefore, A/S and S/A do not impose any insurmountable limitation on the use of the GPS for occultation measurements.

In GPS precision geodesy, "Double Differencing" (DD) is employed to effectively cancel nearly all errors resulting from transmitter clock uncertainty and receiver clock biases. As illustrated in Figure 2 below, the DD technique starts by forming a "DD observable" from the linear combination of 4 observables, each with certain common errors. By differencing observations of a given GPS satellite at 2 receivers, clock errors and S/A dithering for that satellite are canceled. This is referred to as a Single Difference (SD). If SDs are formed for a second GPS satellite and differenced with the first SD, a DD is formed canceling errors common to the receiver clocks. For GPS/MET, a network of ground based receivers, located at precisely known fiducial sites, will be used in conjunction with the data collected from the GPS/MET LEO receiver to implement the DD technique.
Figure 2 Double Differencing Geometry

3.2 Retrieval Methodology

The process described below might be considered the "classical retrieval method". The fundamental principles have evolved over time from the original planetary occultation work conducted at SUCRA and JPL, as described above. Retrieval methods are improving all the time, however, and the GPS/MET team is investigating new techniques. The description of the retrieval methodology which follows was first described in a paper on the GPS occultation technique co-authored by scientists at Lockheed, SUCRA, and JPL.[13]

3.2.1 Compensating for the Ionosphere

To extract information on the neutral atmosphere, propagation delays caused by the ionosphere must be isolated and removed from the signal. Electrons in the ionosphere cause a frequency dependent delay in the phase of received GPS signals. Anticipating the need for ionospheric corrections, GPS planners designed into the system the use of two carrier frequencies, L1 and L2, as previously described. By using dual frequency phase measurements, and knowledge of the inverse square relationship between the group delay and the frequency of each carrier, a simple linear correction can be derived.[14] This correction can be expressed as follows:

\[
TD_{L1} = 1.5336 \times \Delta T \quad (1)
\]

where \(TD_{L1}\) is the ionospheric delay on L1 and \(\Delta T\) is the measurable difference in delay between L1 and L2. The Doppler frequency offset, also affected by the ionosphere, can be modeled with a similar linear correction:

\[
\Delta f_{L1} = 3.529 \times (\Delta f_{L2} - \Delta f_{L1}) \quad (2)
\]

where \((\Delta f_{L2} - \Delta f_{L1})\) is the measurable Doppler difference.[15] Correcting for these ionospheric effects completes the first step in the recovery of meteorological data from the observables.[16]

The method described above provides a simple first order correction for ionospheric effects. In most ground based applications, where the L1 and L2 rays follow substantially identical paths, it is sufficient. And for GPS/MET, it will provide sufficient accuracy for soundings below 30 km. However, for profiles above 30 km, a more sophisticated ionospheric correction scheme is required. To meet the requirement, an advanced technique which takes into account the separation of the L1 and L2 rays is under development by the GPS/MET team.

3.2.2 Recovering Atmospheric Index of Refraction

The fundamental measurement in the radio occultation technique is the time delay of the signal, or resulting phase shift in the signal received from the GPS transmitter.[17] The radio signal propagating from the GPS transmitter to the LEO receiver follows a path through the atmosphere that curves distinctively in response to atmospheric gradients in refractive index. The cumulative effect of the atmosphere on the ray path can be expressed in terms of the total refractive bending angle, \(\alpha\), as shown
in Figure 3 below.

The variation of $\phi$ with experiment geometry can be characterized through use of an "impact parameter", $a$, defined as the perpendicular distance between the center of the planet and the straight line followed by the ray approaching the atmosphere. When combined with a precise knowledge of the geometry (obtained concurrently from other GPS satellites), each sample of phase data (corrected for ionospheric effects) can be converted to the corresponding values for $\phi$ and $a$. This step is straightforward and involves simple geometrical considerations, basic laws of geometrical optics, and relativistic formulas for Doppler shifts.

For an atmosphere with local spherical symmetry (i.e., no significant asymmetric horizontal variations in temperature or moisture), there is a unique relationship between $\phi(a)$ and $\mu(r)$, the atmospheric refractive index as a function of radius ($r$). The refractive index profile $\mu(r)$ is then derived through an Abel transform of the measurements of $\phi(a)$ obtained over the complete occultation, as given in Eq. (3).

$$\mu(r_m) = -\frac{1}{\pi a_m} \int_{a_m}^{\infty} \frac{\phi(a)}{\left(a^2 - a_m^2\right)^{1/2}} da$$

Figure 3 Occultation Geometry

Here $\mu(r_m)$ is the index of refraction of the layer a distance $r_m$ from the center of mass of the planet, and $a_m$ is the value of $a$ for the ray whose radius of closest approach is $r_m$. Application of Eq. (3) layer by layer, starting with the uppermost atmospheric layer and working downward, will provide the index of refraction profile through the atmosphere. This transformation has inherent in it the assumptions that: (1) the atmospheric shells are spherical, and (2) each shell has a uniform index of refraction, i.e., no horizontal variations.

The assumption of spherical symmetry, required for the classical retrieval method, is a limitation which may need to be overcome to achieve the generality desired for an operational system. However, the error introduced by using the assumption of spherical symmetry may not be the dominant error source, and therefore may be acceptable for operational systems. A recent paper by Russian scientists Sokolovskiy and Gorbonov tend to support this possibility.[18] It should be noted that some state-of-the-art ray tracing algorithms developed for seismology do not depend on the assumption of spherical symmetry. We plan to explore the incorporation of these advanced algorithms in our refractivity retrieval approach.
3.2.3 Refractivity to Meteorological Parameters

Classical atmospheric parameters of interest can be derived from the refractive index profile through the following sequence of steps. To simplify the explanation, the process will first be described for the case of dry air. Then, the effect of moisture will be considered.

3.2.3.1 Dry Air Case

First, as the index of refraction, \( \mu \), is close to unity in the terrestrial atmosphere, it is convenient to define the refractivity \( N \):

\[
N = (\mu - 1) \times 10^6
\]  

(4)

For dry air, \( N \) can be expressed as:

\[
N = 77.60 \times (P/T)
\]  

(5)

where \( P \) is the pressure in millibars and \( T \) is the temperature in Kelvins. Furthermore, the equation of state for dry air takes the form:

\[
\rho = 0.3484 \times (P/T)
\]  

(6)

where \( \rho \) is the air density in kg m\(^{-3}\). Equations (5) and (6) show that \( \rho \) is directly proportional to \( N \) for dry air, so that \( \rho(r) \) can be obtained easily from \( \mu(r) \). Next, \( P(r) \) can be obtained from \( \rho(r) \) by integrating the equation of hydrostatic equilibrium:

\[
dP/dh = -g\rho
\]  

(7)

where \( h \) is the height and \( g \) is the acceleration of gravity. Finally, \( T \) can be obtained from \( P \) and \( \rho \) using Eq. (6). In summary, vertical profiles of \( \rho \), \( P \), and \( T \) can be obtained from \( \mu(r) \) in a direct and simple manner.

The total refractive bending angle, \( \alpha \), shown in Figure 3 is greatly exaggerated. For the Earth's atmosphere, the maximum bending angle is on the order of 0.02 radians (1 deg.). To place this in perspective, the phase shift measurements made with the Voyager spacecraft demonstrated that \( \alpha \) could be measured with an accuracy approaching \( 10^{-8} \) radians. With comparable performance from a spaceborne GPS receiver, the refractive bending caused by the terrestrial atmosphere could be resolved to about 1 ppm. It is this type of precision in the radio measurements that leads to the expectation of obtaining high precision vertical profiles of \( N \), \( \rho \), \( P \), and \( T \) in regions of the atmosphere where the air is dry.

3.2.3.2 General Case

The procedure described above must be modified to account for the presence of water vapor. When the
effect of water vapor is included, the expression for the refractivity becomes:

\[
N = 77.60 \times \left( \frac{P}{T} \right) + 3.730 \times 10^5 \times \left( \frac{P_w}{T^2} \right)
\]

(8)

(DRY TERM) (WET TERM)

where \( P_w \) is the vapor pressure of water in millibars. The "dry term" from Eq. (5) has been supplemented by a contribution from water vapor (the "wet term") which can be substantial in the lowest scale height of the atmosphere above the Earth's surface. The moist term also exhibits considerable variation with location and time. The separate contributions to \( N \) by the dry and moist terms cannot be distinguished uniquely through occultation measurements with the current capabilities of the GPS satellites.\[19\] This introduces an ambiguity into the profiles of \( P, P, \) and \( T \); the effects of water vapor at variable and uncertain concentrations are indistinguishable from the effects of background variations in temperature and pressure.

At altitudes above 8-10 km, this ambiguity is not a significant problem as the contribution to the refractive index by water vapor is usually much less than 2%. Similarly, the contribution of moisture to refractive index is negligible throughout the polar atmosphere during winter. In the lower troposphere, the water vapor limitations can be overcome by one of several means, such as use of auxiliary methods for estimation of water vapor content (e.g., through microwave radiometry or ground-based GPS measurements)\[20\] and use of independent temperature measurements at known locations (e.g., radiosondes, aircraft). For example, if the temperature profile in the troposphere was known from model calculations, then moisture profiles could be retrieved from the measurements. This approach will work best in tropical regions where the temperature profiles exhibit relatively small changes, but moisture fields change significantly in space and time. It should be emphasized that \( \mu \) and \( N \) can still be determined accurately regardless of the abundance of water vapor.