GNSS Solutions:

Carrier phase and its measurement for GNSS

"GNSS Solutions" is a regular column featuring questions and answers about technical aspects of GNSS. Readers are invited to send their questions to the columnist, **Dr. Mark Petovello**, Department of Geomatics Engineering, University of Calgary, who will find experts to answer them. His e-mail address can be found with his biography at the conclusion of the column.

What is the carrier phase measurement? How is it generated in GNSS receivers?

Simply put, the carrier phase measurement is a measure of the range between a satellite and receiver expressed in units of cycles of the carrier frequency. This measurement can be made with very high precision (of the order of millimeters), but the whole number of cycles between satellite and receiver is not measurable.

A good analogy to this is to imagine a measuring tape extending from the satellite to the receiver that has numbered markers every one millimeter. Unfortunately, however, the numbering scheme returns to zero with every wavelength (approximately 20 centimeters for GPS L1). This allows us to measure the range very precisely, but with an ambiguity in the number of whole carrier cycles.

A number of subtleties must be considered when generating these measurements in a receiver. To fully appreciate these requires a more detailed look at what we mean precisely when discussing the carrier phase.

All current GNSS satellites transmit radio frequency (RF) signals in the L-band. These signals consist of, at the very least, an RF carrier modulated by a pseudorandom noise (PRN) code. When discussing the phase of a signal it is important to realize that phase is fundamentally a property of sinusoids. Every sinusoid has an amplitude and a phase and can be written in complex notation as $A_{\exp} j(\omega t + \theta)$, where A is the amplitude, ω is the radial frequency, and θ is the phase in radians at $t = 0$.

When the RF carrier is modulated by a PRN code, the resulting signal is no longer a pure sinusoid but can, by Fourier's Theorem, be expressed as a linear combination of sinusoids. The concept of the phase of a combination of sinusoids is less clear than that for a single sinusoid (e.g., how do we define the phase of the sum to two sinusoids?).

For GNSS signal processing we define the signal phase to be the phase of the carrier signal. In other words, this is the phase of the pure sinusoid that would result if the PRN code and any other modulations are "wiped off." In the following discussion, when we mention the signal phase, this is what we shall be referring to.

Consider a sinusoidal signal transmitted from satellite s at time t_{tx} and received at receiver r at time t_{rx}. This signal can be represented as

$$y_r^s(t_{rx}) = A_r^s(t_{rx}) \exp j2\pi\phi_r^s(t_{rx}) \quad (1)$$

where $A_r^s(t)$ is the received amplitude and $\phi_r^s(t)$ is the received phase in cycles. For a plane wave propagating in free space, the phase of the received RF signal is given by

$$\begin{aligned} \phi^s_{r,RF}(t_{rx}) &\equiv f_{RF}\, t_{tx} + \phi^s(0) \pmod 1 \\ &\equiv f_{RF}\, t_{rx} + \phi^s(0) - \frac{r(t_{rx})}{\lambda} \pmod 1 \end{aligned} \quad (2)$$

where f_{RF} is the frequency of the transmitted sinusoid in Hertz (e.g., for GPS L1 f_{RF} 1575.42 MHz), $\phi^s(0)$ is the initial phase at the transmitter, $\lambda = c/f_{RF}$ is the wavelength, and $r(t_{rx}) = (t_{rx} - t_{tx})c$ is the range from satellite to receiver.

The key point to note is that the received carrier phase gives information regarding the range between satellite and receiver. In GNSS applications we would like to extract this information for navigation purposes. Ideally, we want to obtain a measure of the range expressed in units of cycles. Rearranging the previous equation gives us:

$$\frac{r(t_{rx})}{\lambda} = f_{RF}\, t_{rx} + \phi^s(0) - \left[\phi^s_{r,RF}(t_{rx}) + M^s_r(t_{rx})\right] \quad (3)$$

where M^s_r is the true integer number of wavelengths (cycles) between the satellite and the receiver. So, if the receiver can estimate the received RF phase, then three unknowns remain that must be estimated in order to determine the range. These are:

1. The receiver time t_{rx}
2. The initial satellite phase offset $\phi^s(0)$
3. The integer number of cycles between satellite and receiver M^s_r.

In most carrier phase processing applications, the effects of the first two are removed by differencing of observations; differencing between satellites removes errors in the receiver time component, while differencing between receivers removes errors in the satellite phase offset.

So, how does the receiver generate an estimate of the carrier phase observation? The short answer is by integrating the Doppler frequency (hence, why the carrier phase measurement is often called the accumulated Doppler range), but this does not address the original question. So, instead consider the simplified overview of the carrier phase from transmission to intermediate frequency, as given in **Figure 1**.

The incoming signal is first passed through a low noise amplifier (LNA); then it is multiplied by a locally generated sinusoid and filtered, a process known as mixing. The resulting signal is a replica of the RF signal, but shifted down to a lower, intermediate frequency (IF).

The mixer signal is a sinusoid whose phase is given by

$$\phi_{r,Mix}(t) = (f_{RF} - f_{IF})\, t + \phi_{r,Mix}(0) \quad (4)$$

where f_{IF} is the intermediate frequency. The phase of the signal at the mixer output, the IF signal phase, is then given by

$$\phi^s_{r,IF}(t_{rx}) = \phi^s_{r,RF}(t_{rx}) - \phi_{r,Mix}(t_{rx}) \quad (5)$$

$$\equiv \phi^s(0) + f_{IF}\, t_{rx} - \phi_{r,Mix}(0) - \frac{r(t_{rx})}{\lambda} \pmod 1$$

While only a single phase is indicated in Figure 1, the signal at the antenna is the sum of the signals from all satellites in view. At this point in the receiver the IF signal is passed to a number of phase locked loops (PLLs) operating in parallel. **Figure 2** shows a simplified linear model of the PLL. Each PLL tracks the phase of the IF signal for a given satellite.

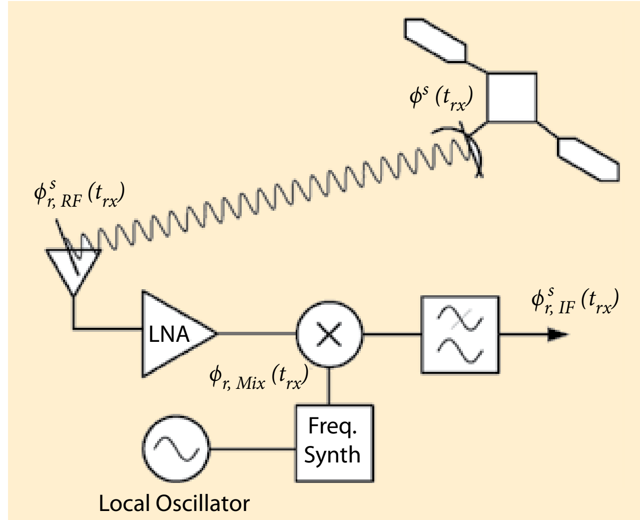


FIGURE 1 Signal phase from transmission through downconversion to intermediate frequency. At each point the signal is assumed to be a sinusoid, the phase of which is indicated in text in the figure.

Once the PLL is locked to the incoming signal, $\phi^s_{r,NCO} \equiv \phi^s_{r,IF}$ (mod 1). Comparing Equations (3) and (5), it can be seen that only the term $f_{IF} t_{rx}$ is known by the receiver, and hence the carrier phase measurement, denoted ϕ^s_r,

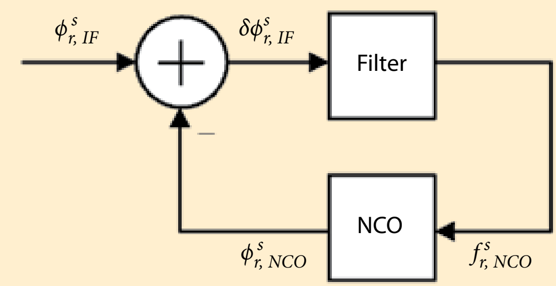


FIGURE 2 Linear model of a phase locked loop (PLL). A local replica sinusoid with phase $\phi^s_{r,NCO}$ is generated by a numerically controlled oscillator (NCO). The difference between this local replica and the IF phase is filtered to generate a command signal for the NCO $f^s_{r,NCO}$, which drives the error signal $\delta\phi^s_{r,IF}$ to zero.

can be generated as follows

$$\phi^s_r(t_{rx}) = f_{IF}t_{rx} - \phi^s_{r,NCO}(t_{rx}) \equiv \frac{r(t_{rx})}{\lambda} + \phi_{r,Mix}(0) - \phi^s(0) \pmod 1 \tag{6}$$

Note that, rather than being a measure of the true range, the carrier phase measurement also includes terms due to phase offsets in the receiver and the satellite. As such, this measurement can be considered to be an ambiguous measure of the pseudorange, expressed in cycles of the carrier.

A number of problems arise with generating the carrier phase measurement in this way, however. First, note that we can consider the NCO phase to consist of two components, one due to the intermediate frequency and another that is just due to the range between satellite and receiver.

Consider the simple case of an IF of 0 Hz. In this case the NCO phase consists only of the range component, and the carrier phase measurement can be obtained directly as the negative of the NCO phase. Unfortunately the integer number of cycles between satellite and receiver $M^s_r(t_{rx})$ is not observable in the PLL.

To generate useable phase measurements, the receiver phase observations must maintain a constant integer number of cycles offset from the true carrier phase. To do this, both integer and fractional components of the NCO phase must be stored. To see this more clearly, consider the situation where phase tracking of the signal from satellite s begins at time t_s, at which point the phase estimate is initialized to zero (i.e., both fractional and integer components are zero):

$$\phi^s_r(t_s) = 0 \tag{7}$$

At some later time t_s, phase lock is achieved, and the fractional phase of the local replica signal is a good estimate of the fractional phase of the down-converted satellite signal. However, the integer component of the phase is still offset from the true integer component by some arbitrary unknown number of cycles. This offset is referred to as the *carrier phase ambiguity*.

Let $L^s_r(t_{rx})$ denote the integer component of the NCO phase, then the phase observation is given by

$$\phi_r^s(t_{rx}) = \left(\frac{r(t_{rx})}{\lambda} + \phi_{r,Mix}(0) - \phi^s(0)\right) \bmod 1 + L_r^s(t_{rx}) \tag{8}$$
$$= \frac{r(t_{rx})}{\lambda} + \phi_{r,Mix}(0) - \phi^s(0) + \left[L_r^s(t_{rx}) - K_r^s(t_{rx})\right]$$
$$= \frac{r(t_{rx})}{\lambda} + \phi_{r,Mix}(0) - \phi^s(0) + N_r^s(t_{rx})$$

where $K_r^s(t_{rx})$ is the integer component of the term in parentheses on the first line of Equation (8), and $N_r^s(t_{rx}) = L_r^s(t_{rx}) - K_r^s(t_{rx})$ is known as the integer ambiguity. Provided that continuous phase tracking is maintained, this term should be constant.

So, for the case of zero IF, the only consideration for generating useable phase measurements is to ensure that the integer ambiguity remains constant, that is, if the range increases by one cycle, the integer component of the NCO phase, $L_r^s(t_{rx})$, also increments by one cycle. This can be accomplished by simply integrating the NCO frequency and incrementing the number of integer cycles.

If the IF is non-zero, there is an extra consideration to be made. Now the NCO phase consists of two components:

1. The IF phase: $\phi^s_{r,\,NCO,\,IF}$. This component of the NCO phase should ideally match the IF phase term at the mixer output, $f_{IF}t_{rx}$.
2. The negative of the phase measurement: $-\phi_r^s$

The total NCO phase is then given by:

$$\phi^s_{r,\,NCO}(t_{rx}) = \phi^s_{r,\,NCO,\,IF}(t_{rx}) - \phi_r^s(t_{rx}) \tag{9}$$

The integer component of the IF phase is irrelevant and can be discarded. Again, consider that phase tracking of the signal from satellite s commences at time t_s; so, the phase measurement component of the NCO phase is again initialized to zero.

Care must be taken as to how the IF phase component is initialized. Recall that the PLL drives the difference between the incoming IF phase and the total NCO phase to zero; thus, once phase lock is achieved the total NCO phase will be given by

$$\phi^s_{r,NCO}(t_{rx}) = \phi^s_{r,NCO,IF}(t_s) + (t_{rx} - t_s) f_{IF} - \phi_r^s(t_{rx}) \tag{10}$$
$$\equiv \phi^s(0) + f_{IF} t_{rx} - \phi_{r,\,Mix}(0) - \frac{r(t_{rx})}{\lambda} \pmod 1$$
$$\therefore \phi_r^s(t_{rx}) = \frac{r(t_{rx})}{\lambda} + \phi_{r,\,Mix}(0) + t_s f_{IF} - \phi^s_{r,\,NCO,\,IF}(t_s) - \phi^s(0) + N_r^s(t_{rx})$$

Comparing Equations (8) and (10), two extra constant terms appear in the phase measurement in Equation (10). These terms (third and fourth) are a function both of the time at which phase tracking begins, t_s, and the initial value of the IF component of the NCO phase, $\phi^s_{r,\,NCO,\,IF}(t_s)$.

This phase bias term is therefore different for each satellite, which could cause significant problems in processing these measurements. However, a receiver designer is free

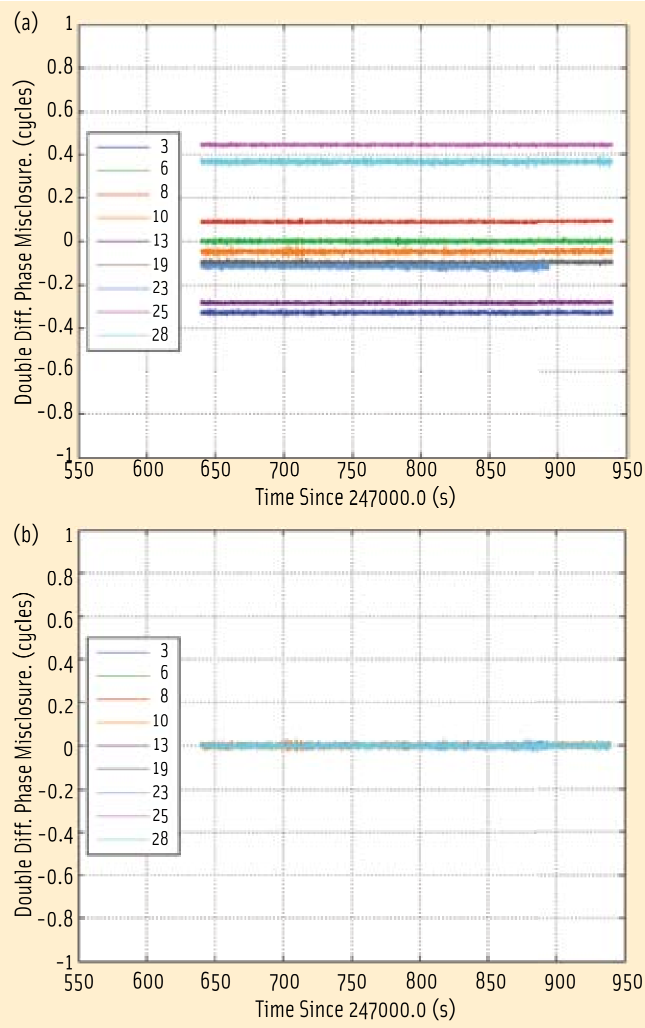


FIGURE 3 **Double difference phase misclosures for a zero-baseline test: a) IF phase initialized to 0, b) global IF phase accumulator used.**

to choose the initial value of the NCO phase and thus can ensure that this bias is the same for all satellites.

One simple way to do this is to mix the IF signal down to zero, thereby reverting to the zero IF case. An alternative is to store a globally accessible IF phase accumulator in the receiver. This can be arbitrarily initialized when the receiver is turned on, and its phase is updated continuously during receiver operation. Whenever a PLL starts to track a signal, the IF component of its NCO phase is initialized to that of the global IF phase accumulator.

The initial value of the IF component for satellite s is therefore given by

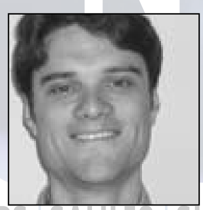
$$\phi^{s}_{r,\,NCO,IF}\left(t_s\right)=\left(t_s-t_0\right)f_{IF} \tag{11}$$

where t_0 is the time at which the receiver is turned on. Inserting this into Equation (10), we see that the extra phase bias term is now given by $t_0 f_{IF}$, which is the same for all satellites tracked. This term can therefore be absorbed into the initial mixer phase term in Equation (10).

To demonstrate the foregoing explanation, raw IF data was collected in a zero-baseline configuration with a non-zero value for f_{IF}. The carrier phase measurements were initialized in two ways: a) the IF phase was initialized to zero, and b) the IF phase was initialized using the global IF phase accumulator. The double difference misclosures for each case are plotted in **Figure 3**. The satellite dependent phase biases are clearly visible in case a), shown in the figure's upper panel, while there are no such biases in case b), the lower panel.

In summary, the carrier phase measurement is a highly precise measure of the pseudorange between satellite and receiver, the generation of useable carrier phase measurements in a receiver requires a phase locked loop, and the receiver designer must take care to ensure that 1) the integer ambiguity term is constant, and 2) the initial phase is chosen so as to ensure a common phase bias in all tracking channels.

CILLIAN O'DRISCOLL



DR. CILLIAN O'DRISCOLL is a senior research engineer at the University of Calgary. His research interests are in all areas of GNSS signal processing.

Correction

On page 22 of the June issue's GNSS Solutions column, the final term at the end of the first paragraph is missing a parenthesis. It should be: Sinc(x) = $sin(\pi x)/(\pi x)$.

On page 24, the last sentence in the fourth paragraph should read: "In this region the approximation is quite good and degrades only in the points where the values given by (3) are of the same order of the terms of the integrated double frequency." Also, in the seventh paragraph of the same page, the term $\Delta\tau > 1$ has an inverted symbol. The complete sentence should read: "Moreover, we can observe that for small values of $\Delta\tau$ (that is, $\Delta\tau < 1$ chip duration) Equation (3) also gives a good approximation of the AF, as shown in **Figure 4**."

Finally, the title for column editor Mark Petovello should have been "Associate Professor," following his promotion in April. IG

Mark Petovello is an Associate Professor in the Department of Geomatics Engineering at the University of Calgary. He has been actively involved in many aspects of positioning and navigation since 1997 including GNSS algorithm development, inertial navigation, sensor integration, and software development.

Email: mark.petovello@ucalgary.ca