Cambridge University Press
978-0-521-88677-2 - Phase Noise and Frequency Stability in Oscillators
Enrico Rubiola
Excerpt
More information

# 1 Phase noise and frequency stability

In *theoretical physics*, the word "oscillator" refers to a physical object or quantity oscillating sinusoidally – or at least periodically – for a long time, ideally forever, without losing its initial energy. An example of an oscillator is the classical atom, where the electrons rotate steadily around the nucleus. Conversely, in *experimental science* the word "oscillator" stands for an artifact that delivers a periodic signal, powered by a suitable source of energy. In this book we will always be referring to the artifact. Examples are the hydrogen maser, the magnetron of a microwave oven, and the swing wheel of a luxury wrist watch. Strictly, a "clock" consists of an oscillator followed by a gearbox that counts the number of cycles and the fraction thereof. In digital electronics, the oscillator that sets the timing of a system is also referred to as the clock. Sometimes the term "atomic clock" is improperly used to mean an oscillator stabilized to an atomic transition, because this type of oscillator is most often used for timekeeping.

A large part of this book is about the "precision"[1] of the oscillator frequency and about the mechanisms of frequency and phase fluctuations. Before tackling the main subject, we have to go through the technical language behind the word "precision," and present some elementary mathematical tools used to describe the frequency and phase fluctuations.

## 1.1 Narrow-band signals

The ideal oscillator delivers a signal

$$v(t) = V_0 \cos(\omega_0 t + \varphi) \qquad \text{(pure sinusoid)}\,, \tag{1.1}$$

where $V_0 = \sqrt{2}\, V_{\text{rms}}$ is the peak amplitude, $\omega_0 = 2\pi \nu_0$ is the angular frequency,[2] and $\varphi$ is a constant that we can set to zero. Let us start by reviewing some useful representations associated with (1.1).

A popular way of representing a noise-free sinusoid $v(t)$ in Cartesian coordinates is the *phasor*, also called the *Fresnel vector*. The phasor is the complex number $\mathbf{V} = A + jB$ associated with $v(t)$ after factoring out the $\omega_0 t$ oscillation. The absolute value

[1] Here, the word "precision" is not yet used as a technical term.
[2] The symbol $\omega$ is used for the angular frequency. Whenever there is no ambiguity, we will omit the adjective "angular" and give the numerical value in Hz, which of course refers to $\nu_0 = \omega/(2\pi)$.

$|\mathbf{V}|$ is equal to the rms value of $v(t)$, and the phase $\arg \mathbf{V}$ is equal to $\varphi$. The phase reference is set by $\cos \omega_0 t$. Alternatively, the phasor is obtained by expanding $v(t)$ as $V_0 (\cos \omega_0 t \cos \varphi - \sin \omega_0 t \sin \varphi)$. Then, the real part is identified with the rms value of the $\cos \omega_0 t$ component and the imaginary part with the rms value of the $- \sin \omega_0 t$ component. Thus, the ideal signal (1.1) may be represented as the phasor

$$\begin{aligned} \mathbf{V} &= \frac{V_0}{\sqrt{2}}\, e^{j\varphi} \\ \mathbf{V} &= \frac{V_0}{\sqrt{2}}\, (\cos \varphi + j \sin \varphi) \end{aligned} \qquad \text{(phasor)}\,. \tag{1.2}$$

A more powerful tool is the *analytic signal* $z(t)$ associated with $v(t)$, also called the *pre-envelope* and formally defined as

$$z(t) = v(t) + j\hat{v}(t) \qquad \text{(analytic signal)}\,, \tag{1.3}$$

where $\hat{v}(t)$ is the Hilbert transform of $v(t)$, i.e. $v(t)$ shifted by 90°. The analytic signal is most often used to represent narrowband signals, i.e. signals whose power is clustered in a narrow band centered at the frequency $\omega_0$. However, it is not formally required that the bandwidth be narrow, nor that the power be centered at $\omega_0$, and not even that $\omega_0$ be contained in the power bandwidth. Amplitude and phase can be (slowly) time-varying signals.

The analytic signal $z(t)$ is obtained from $v(t)$ by deleting the negative-frequency side of the spectrum and multiplying the positive-frequency side by a factor 2. Alternatively, $z(t)$ can be obtained using any of the following replacements:

$$v(t) = \frac{V_0}{\sqrt{2}} \cos(\omega_0 t + \varphi) \quad \Rightarrow \quad \begin{cases} z(t) = V(t) e^{j[\omega_0 t + \varphi(t)]} \\ z(t) = V(t) e^{j\varphi(t)}\, e^{j\omega_0 t} \\ z(t) = V(t) (\cos \varphi + j \sin \varphi)\, e^{j\omega_0 t}\,. \end{cases} \tag{1.4}$$

The analytic signal has two relevant properties.

1. A phase shift $\theta$ applied to $v(t)$ is represented as $z(t)$ multiplied by $e^{j\theta}$.
2. Since the power associated with negative frequencies is zero, the total signal power can be calculated using the positive frequencies only.

The *complex envelope* of $z(t)$, also referred to as the *low-pass* process associated with $z(t)$, is obtained by deleting the complex oscillation $e^{j\omega_0 t}$ in the analytic signal. The complex envelope is the natural extension of the phasor and is used when the amplitude and phase are allowed to vary with time:

$$\mathbf{V} = \frac{V_0}{\sqrt{2}}\, e^{j\varphi} \qquad \Longleftrightarrow \qquad \tilde{v}(t) = V(t)\, e^{j\varphi(t)}\,. \tag{1.5}$$

Strictly speaking, the phasor refers to a pure sinusoid. Yet the terms "phasor" and "time-varying phasor" are sometimes used in lieu of the term "complex envelope."

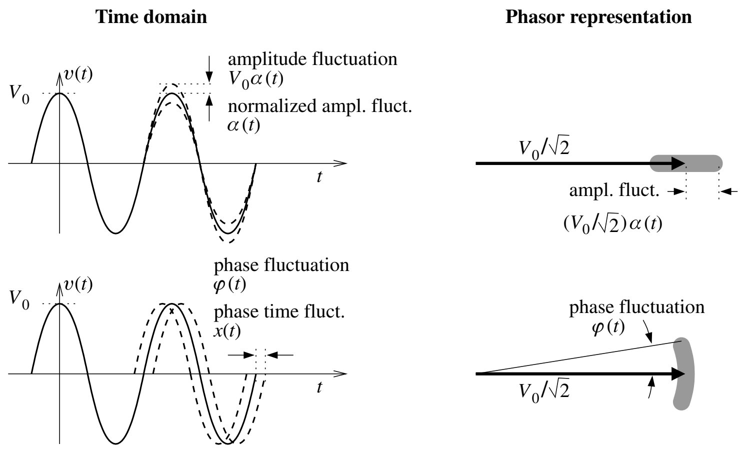


**Figure 1.1** Amplitude and phase noise: $V_0$ is in volts, $\alpha(t)$ is non-dimensional, $\varphi(t)$ is in radians, and $x(t)$ is in seconds.

### 1.1.1 The clock signal

In the real world, an oscillator signal fluctuates in amplitude and phase. We introduce the quasi-perfect sinusoidal *clock signal* (Fig. 1.1)

$$v(t) = V_0[1 + \alpha(t)] \cos[\omega_0 t + \varphi(t)] \, , \qquad |\alpha(t)| \ll 1 \, , \quad |\varphi(t)| \ll 1 \, . \tag{1.6}$$

The term "clock signal" emphasizes the fact that the cycles of $v(t)$, and fractions thereof, can be counted by suitable circuits, so that $v(t)$ sets a time scale. When talking about clocks, we assume that (1.6) has a high signal-to-noise ratio. Hence we note the following.

- The peak amplitude $V_0$ of (1.1) is replaced by the envelope $V_0[1 + \alpha(t)]$, where $\alpha(t)$ is the random fractional amplitude.[3] The assumption

$$|\alpha(t)| \ll 1 \tag{1.7}$$

  reflects the fact that actual oscillators have small amplitude fluctuations. Values $|\alpha(t)| \in (10^{-3}, 10^{-6})$ are common in electronic oscillators.
- The constant phase $\varphi$ of (1.1) is replaced by the random phase $\varphi(t)$, which originates the clock error. In most cases, we can assume that

$$|\varphi(t)| \ll 1 \, . \tag{1.8}$$

A slowly varying phase is often referred to as *drift*. Observing a clock in the long term, the assumption $|\varphi(t)| \ll 1$ is no longer true. Yet it is possible to divide the carrier frequency by a suitably large rational number. The phase scales down accordingly, so that the condition $|\varphi(t)| \ll 1$ is obtained at low frequencies.

[3] The symbol $\epsilon(t)$ is often used in the literature instead of $\alpha(t)$.

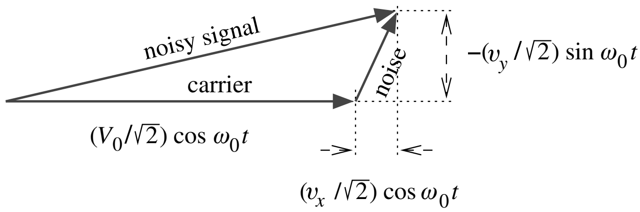


**Figure 1.2** Phasor representation of a noisy sinusoid.

The clock signal can be rewritten in Cartesian coordinates by separating the $\cos\omega_0 t$ and $-\sin\omega_0 t$ components[4] (Fig. 1.2):

$$v(t) = V_0\cos\omega_0 t + v_x(t)\cos\omega_0 t - v_y(t)\sin\omega_0 t\ . \tag{1.9}$$

The signals $v_x(t)$ and $v_y(t)$ are called the *in-phase* and *quadrature* components of the noise, respectively.

The polar representation (1.6) and the Cartesian representation (1.9) are connected by

$$\alpha(t) = \sqrt{\left[1+\frac{v_x(t)}{V_0}\right]^2 + \left[\frac{v_y(t)}{V_0}\right]^2} - 1 \qquad \text{(fractional amplitude)}\ , \tag{1.10}$$

$$\varphi(t) = \arctan\frac{v_y(t)}{V_0 + v_x(t)} \qquad \text{(phase)}\ . \tag{1.11}$$

Equation (1.10) is the Pythagorean theorem written in terms of the real component $(V_0 + v_x)/\sqrt{2}$ and the imaginary component $v_y/\sqrt{2}$. Equation (1.11) is arctan $\Im/\Re$, i.e. the arctangent of the imaginary-to-real ratio. A problem with (1.11) is that the arctangent returns the principal value, i.e. the value defined in $(-\pi/2, \pi/2)$, or in $(-\pi, \pi)$ using a two-argument arctangent. The cycles accumulated, if any, are to be counted separately. In low-noise conditions, it holds that, for $|v_x/V_0| \ll 1$ and $|v_y/V_0| \ll 1$,

$$\alpha(t) = \frac{v_x(t)}{V_0} \qquad \text{(fractional amplitude)} \tag{1.12}$$

$$\varphi(t) = \frac{v_y(t)}{V_0} \qquad \text{(phase)}\ . \tag{1.13}$$

The spectrum of a pure sinusoid such as (1.1) is an ideally thin line at $\omega_0$, mathematically described as a Dirac delta function $\delta(\omega - \omega_0)$. Noise broadens the spectrum: the clock signal (1.6) looks like a line of bandwidth twice that of $\alpha(t)$ and $\varphi(t)$ if the signal-to-noise ratio is high. In the case of a low signal-to-noise ratio, $\varphi(t)$ yields a linewidth larger than twice the bandwidth.

Finally, the random phase $\varphi(t)$ does not contribute to the signal power. The instantaneous power is $P(t) = v^2(t)/R_0$. In low-noise conditions the power, averaged over

[4] The form $x(t)\cos\omega_0 t - y(t)\sin\omega_0 t$ is preferable for a signal in Cartesian coordinates, but in this chapter $x$ and $y$ are used for other relevant quantities.

a time $T_m$ longer than the oscillation period yet shorter than the time scale $T_\alpha$ of the amplitude fluctuations, is

$$\overline{P(t)} = \frac{V_0^2}{2R_0}\left[1 + 2\alpha(t)\right] , \qquad |\alpha(t)| \ll 1, \quad 2\pi/\omega_0 \ll T_m \ll T_\alpha \,. \tag{1.14}$$

---

**Example 1.1.** Let us estimate the error accumulated in 1 year by a clock based on a 10 MHz oscillator accurate to within $10^{-10}$. The maximum clock error is $T_e = (\Delta\omega/\omega)T_{meas}$; thus $T_e = 10^{-10} \times 3.16 \times 10^7$ s $= 3.2$ ms in one year. The oscillation period is $T_c = 2\pi/\omega_0 = 10^{-7}$ s. Hence the clock error accumulated in one year is $n = T_e/T_c = 3.16 \times 10^4$ cycles of the 10 MHz carrier; thus $\varphi = 2\pi n = 2 \times 10^5$ rad.

---

## 1.2 Physical quantities of interest

Traditionally, physicists use the symbol $\nu$ for the frequency while electrical engineers prefer $f$. It is unfortunate that in the domain of time-and-frequency metrology the notation is sometimes unclear or difficult to understand because both $\nu$ and $f$ are found in the same context. When I came to metrology in the early 1980s with a background in electronics and telecommunications, it took me a long time to get used to this unnecessary complication. The early articles about frequency stability use $\nu$ for fixed frequencies, such as a carrier or a beat note, and $f$ for the Fourier frequency, i.e. the variable of spectral analysis. Other articles use $\nu$ in the carrier signal $\cos 2\pi\nu_0 t$ and $f$ for the spectral analysis of the low-pass fluctuations ($\alpha$, $\varphi$, etc.), considering the carrier and the low-pass fluctuations as nearly separate worlds. Of course, these two distinctions between $\nu$ and $f$ are similar, so it may be difficult to decide whether a frequency should be $\nu$ or $f$. Additional confusion arises from the fact that fluctuations even smaller than $10^{-16}$ are sometimes measured;[5] no frequency can be taken as constant. In the end, it is recommended that the exact meaning of a frequency symbol in a particular equation is always checked.

Another point is that the oscillator instability can be described as a phase fluctuation or as a frequency fluctuation. The first choice is made in the definition of the clock signal (1.6), here repeated:

$$v(t) = V_0[1 + \alpha(t)]\cos[\omega_0 t + \varphi(t)] \,.$$

In this representation, it is implicitly assumed that $\omega_0$ is the best estimate of the oscillator frequency, so that $\omega_0 t$ describes the oscillation and $\varphi(t)$ describes its phase fluctuation. This approach is suitable for short-term measurements, where the oscillator stability is sufficient for $\varphi(t)$ to stay in the interval $(-\pi, \pi)$. In the longer term, the oscillator

[5] It is instructive to relate this small value to the internal computer representation of numbers. Since the IEEE standard "double precision" format has a 15 digit mantissa, the value $10^{-16}$ is one order of magnitude smaller than the roundoff error.

ends up drifting more than a half-cycle of the carrier frequency and the phase $\varphi(t)$ becomes ambiguous. In such cases, we may prefer to characterize the oscillation using the frequency fluctuations. Then the clock signal is written as

$$v(t) = V_0[1 + \alpha(t)] \cos\left[\omega_0 t + \int (\Delta\omega)(t)\, dt\right] \qquad \text{(clock signal)}\,, \tag{1.15}$$

where

$$(\Delta\omega)(t) = \dot{\varphi}(t) \qquad \text{(angular-frequency fluctuation)} \tag{1.16}$$

is the angular frequency fluctuation.

Additionally, it is often useful to normalize $\varphi(t)$ in order to transform the phase noise into time fluctuations, expressed in seconds (see below), and to normalize the oscillator frequency.

The definitions summarized below are aimed at giving straightforward access to the general literature on phase noise and frequency stability; Fig. 1.3 relates the quantities described to a typical experimental setup.

$\varphi(t)$ represents the phase noise, i.e. the random phase fluctuation defined by (1.6).

$\alpha(t)$ represents the fractional-amplitude noise (for short, "amplitude noise"), i.e. the random amplitude fluctuation defined by (1.6).

$\nu$ is used for the carrier frequency, either radio, microwave, or optical, and also for the beat note between two carriers. The symbol $\nu$ can be a variable, as on the frequency axis of a spectrum analyzer, or a constant, as in the nearly constant frequency of an oscillator. We can also find $\nu(t)$ used in the same way as the quantity $(\Delta\nu)(t)$ introduced below.
Example: let $\nu_1 = 10$ GHz and $\nu_2 = 10.24$ GHz be the frequency of two oscillators. On the spectrum analyzer, we see two lines at $\nu = \nu_1$ and $\nu = \nu_2$. After mixing, the beat frequency is $\nu_b = \nu_2 - \nu_1 = 240$ MHz.

$\Delta\nu = \nu - \nu_0$ is the difference between the actual frequency $\nu$ and a reference value $\nu_0$. The latter can be the nominal frequency or a reference value close to $\nu$.

$(\Delta\nu)(t)$ represents the instantaneous frequency fluctuation (or noise). This implies the assumption of slow modulation with a high modulation index, so that the signal can be approximated by a slow-swinging carrier. This means that the carrier and sidebands degenerate into a single Dirac $\delta$ function that tracks the modulation. In most practical cases, $(\Delta\nu)(t)$ should be regarded as the fluctuating output of a frequency comparator, after discarding the dc component.

$f$ is used for frequency in the spectral analysis of low-pass processes, close to dc, after detection. Thus, $f$ is used in connection with $\alpha(t)$, $\varphi(t)$, $v_x(t)$, $v_y(t)$, etc. The symbol $f$ can refer to a variable, as on the frequency axis of a FFT[6] analyzer, or to a constant.

[6] There is no theoretical need to use a FFT (fast Fourier transform) analyzer to measure a near-dc process. However, this is the type of spectrum analyzer used in virtually all cases.

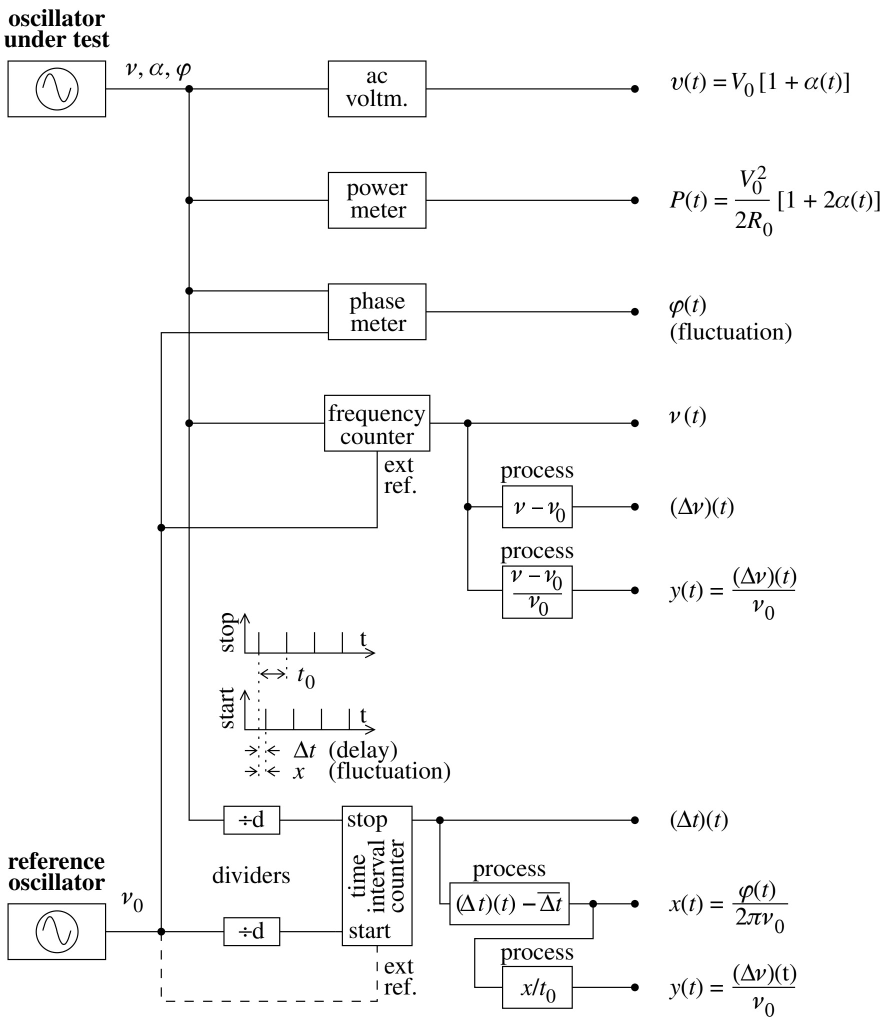


**Figure 1.3** Simplified diagram of oscillator noise measurements in the time domain, illustrating the main physical quantities of interest in time-and-frequency metrology.

Example: inspecting the phase noise of a microwave amplifier in some specific conditions, we find white noise at high $f$ and flicker noise below the corner frequency $f_c$.

$\omega$ is the angular frequency. The unit associated with $\omega$ is rad/s. In this book, $\omega$ is a shorthand for either $2\pi\nu$ or $2\pi f$. We also use $\Delta\omega$ and $(\Delta\omega)(t)$. The word "angular" is often omitted, and numerical values are given in Hz, which of course refers to $\omega/(2\pi)$.

$\Omega$ is used instead of $\omega$ in some special cases, as in Chapter 5, where we need to represent several angular frequencies at the same time. Preferentially, $\Omega$ refers to low-pass phenomena.

$x(t)$ is the phase-time fluctuation, that is, the random phase fluctuation $\varphi(t)$ converted into time, and measured in seconds:

$$x(t) = \frac{\varphi(t)}{\omega_0} = \frac{\varphi(t)}{2\pi\nu_0} \qquad \text{(phase-time fluctuation)}\,. \tag{1.17}$$

Here $\nu_0$ is either the nominal or the estimated frequency.
Interestingly, $x(t)$ does not become ambiguous when $\varphi(t)$ exceeds half a cycle of the carrier (Fig. 1.3) because it allows the accumulation of phase cycles.

$y(t)$ is the fractional-frequency fluctuation, i.e. the instantaneous frequency fluctuation normalized to the carrier frequency $\nu_0$. The quantity $y(t)$ is dimensionless.

$$\begin{aligned} y(t) = \dot{x}(t) &= \frac{\dot{\varphi}(t)}{\omega_0} \\ &= \frac{(\Delta\omega)(t)}{\omega_0} = \frac{(\Delta\nu)(t)}{\nu_0} \qquad \text{(fractional-frequency fluctuation)}\,. \end{aligned} \tag{1.18}$$

The output of digital instruments is a stream of sampled values, denoted by an integer subscript. Thus, $x_k$ is $x(t)$ sampled at the time $t = k\tau$.

Finally, a number of relationships are written in their usual form, found in most textbooks. Therefore, it is inevitable that some of the above symbols are also used in expressions such as "let $f(x)$ be a function . . . ," etc.

Three frequencies have a special rôle all through this book and are used extensively in Chapters 4 and 5; thus they deserve to be mentioned here.

$\omega_0$ and $\nu_0$ are the oscillator frequency and angular frequency, i.e. those of the carrier.
$\omega_n$ and $\nu_n$ are the natural angular frequency and natural frequency of a resonator.
$\omega_p$ and $\nu_p$ are the free-decay pseudofrequency and angular pseudofrequency of a resonator. It holds that $\omega_p \lesssim \omega_0$.

In most oscillators, the oscillation frequency $\omega_0$ is determined by the natural frequency $\omega_n$ of a resonator. However, $\omega_0$ differs slightly from $\omega_n$ because of feedback.

### 1.2.1 ⋆ Frequency synthesis

In a large number of applications the oscillator is the reference of a frequency synthesizer. A noise-free synthesizer can be regarded as a gearbox that multiplies the input frequency $\omega_i$ by a rational number $\mathcal{N}/\mathcal{D}$ and outputs a frequency

$$\omega_o = \frac{\mathcal{N}}{\mathcal{D}}\,\omega_i \qquad \left(\nu_o = \frac{\mathcal{N}}{\mathcal{D}}\,\nu_i\right)\,. \tag{1.19}$$

In the presence of small fluctuations, the input fluctuations propagate to the synthesizer output with the same $\mathcal{N}/\mathcal{D}$ law, that is,

$$(\Delta\omega_o)(t) = \frac{\mathcal{N}}{\mathcal{D}}(\Delta\omega_i)(t) \qquad \left((\Delta\nu_o)(t) = \frac{\mathcal{N}}{\mathcal{D}}(\Delta\nu_i)(t)\right)\,, \tag{1.20}$$

$$\varphi_o(t) = \frac{\mathcal{N}}{\mathcal{D}}\,\varphi_i(t)\,. \tag{1.21}$$

A time lag can be present from input to output if the synthesizer includes a phase-locked loop (PLL). However, the fractional frequency fluctuation and the phase-time fluctuation observed at the output are equal to the input fluctuations:

$$y_{\text{o}}(t) = y_{\text{i}}(t)\ , \tag{1.22}$$

$$x_{\text{o}}(t) = x_{\text{i}}(t)\ . \tag{1.23}$$

Finally, we notice that there is no general law for the propagatation of the amplitude fluctuation $\alpha(t)$ through a synthesis chain. The reason is that the synthesis needs strong nonlinearity, hence the amplitude is saturated.

### Large phase noise

The above rules hold when the phase noise is low. In large-phase-noise conditions, the synthesizer's behavior is governed by the energy conservation law. The easiest way to understand this is to write the output signal in the analytic form $z(t) = V_{\text{rms}}\, e^{j\omega_0 t}\, e^{j\varphi(t)}$. The phase-noise term $e^{j\varphi(t)}$ spreads the power into the noise sidebands, yet without changing the total power because $|e^{j\varphi(t)}| = 1$. If the output phase noise exceeds some 2 radians, the sidebands sink most of the power and the carrier power drops abruptly. This phenomenon is referred to as *carrier collapse*. As a consequence, extremely high spectral purity is needed when the multiplication ratio is high, for example in the synthesis of THz or optical signals from electronic oscillators.

Mathematically, carrier collapse is a consequence of the application of the Angers–Jacobi expansion

$$e^{jz\cos\alpha} = \sum_{n=-\infty}^{\infty} J_{\text{n}}(z)\, e^{jn\alpha} \tag{1.24}$$

to the angular modulation, according to which the carrier amplitude is dominated by the Bessel function $J_0(z)$. The function $J_0(z)$ nulls at $z \simeq 2.405$. Further consideration of this phenomenon is beyond our scope, however.

## 1.3 Elements of statistics

In the previous sections, we expressed the phase noise and amplitude noise in terms of simple time-dependent functions, denoted by $\alpha(t)$ and $\varphi(t)$. In reality, to address the nature of noise properly some statistical tools are necessary, and the concept of a random process needs to be introduced. A few definitions are given below to establish the vocabulary. The reader is encouraged to study the subject using appropriate references, among which I prefer [32, 74].

### 1.3.1 Basic definitions

*Random or stochastic process*

A *random process* is defined through a random experiment **e** that associates a time-domain function $x_e(t)$ with each outcome $e$. The specification of such an experiment, together with a probability law, defines a random process $\mathbf{x}(t)$. Each random process has an infinite number of *realizations*, which form an *ensemble*. A realization, also called a *sample function*, is a time-domain signal $x_\mathrm{e}(t)$. For short, the subscript $e$ is dropped whenever there is no ambiguity or no need to refer to a specific outcome $e$.

A random process and its associated ensemble are powerful mathematical concepts, but they are not directly accessible to the experimentalist, who can only measure a finite number of realizations.

*Mean, time average, and expectation*

In the measurement of random processes (subsection 1.3.2) we use simultaneously three types of "average," the simple mean, the time average, and the mathematical expectation. Hence, for clarity we need different notation for these.

Given a series of $N$ data $x_i$, the simple *mean* of $x$ is denoted by angle brackets:

$$\langle x \rangle_N = \frac{1}{N} \sum_{i=1}^{N} x_i \qquad \text{(mean)}\ . \tag{1.25}$$

The simple mean is often used to average the output stream of an instrument. The quantity $x$ is unspecified. For example, we can average in this way a series of numbers, a series of spectra, etc.

The *time average* of $x$ is denoted by an overbar:

$$\overline{x} = \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, dt \qquad \text{(time average)} \tag{1.26}$$

In the case of causal systems, where the response starts at $t = 0$, the integration limits can be changed from $-T/2$ and $T/2$ to 0 and $T$. In most cases, the readout of an instrument is of the form (1.26). This means that the input quantity $x$ is averaged uniformly over the time $T$.

More generally, the time average includes a weight function $w(t)$:

$$\overline{x} = \int_{-\infty}^{\infty} x(t)\, w(t)\, dt \qquad \text{(weighted time average)} \tag{1.27}$$

with

$$\int_{-\infty}^{\infty} w(t)\, dt = 1\ .$$

In theoretical discussions the definition (1.27) is generally adopted as the *definition of the measure of* $x$. The readout of sophisticated instruments can be of this type.