Chapter 4

Statistical Analysis of the Quantizer Output

4.1 PDF AND CF OF THE QUANTIZER OUTPUT

Signal quantization is a nonlinear operation, and is therefore difficult to analyze directly. Instead of devoting attention to the signal being quantized, we shall consider its probability density function. It will be seen that the PDF of the quantized variable may be obtained by strictly linear operations performed on the PDF of the non-quantized variable. So, although quantization acts nonlinearly on signals, it acts linearly on their probability densities.

The analysis that follows will be developed for sampled signals.[1] High-order random processes are readily described in sampled form. If the samples of a random variable x are all independent of each other, a first-order probability density $f_x(x)$ completely describes the statistics of the samples. The characteristic function or CF is the Fourier transform of the PDF (see Eq. (2.17)),

$$\Phi_x(u) = \int_{-\infty}^{\infty} f_x(x) e^{jux} \, \mathrm{d}x \, . \tag{4.1}$$

A quantizer input variable x may take on a continuum of amplitudes, while the quantizer output x' assumes only discrete amplitudes. The probability density of the output $f_{x'}(x')$ consists of a series of Dirac impulses that are uniformly spaced along the amplitude axis, each one centered in a quantization box.

Fig. 4.1 shows how the output PDF is derived from that of the input. Any input event (signal amplitude) occurring within a given quantization box is "reported" at the quantizer output as being at the center of that box. The quantizer is shown in Fig. 4.1(a). The input PDF $f_x(x)$ is sketched in Fig. 4.1(b). The quantization boxes

[1]The basic ideas and some figures of the following sections were published in, and are reprinted here with permission from Widrow, B., Kollár, I. and Liu, M.-C., "Statistical theory of quantization," *IEEE Transactions on Instrumentation and Measurement* 45(6): 35361.

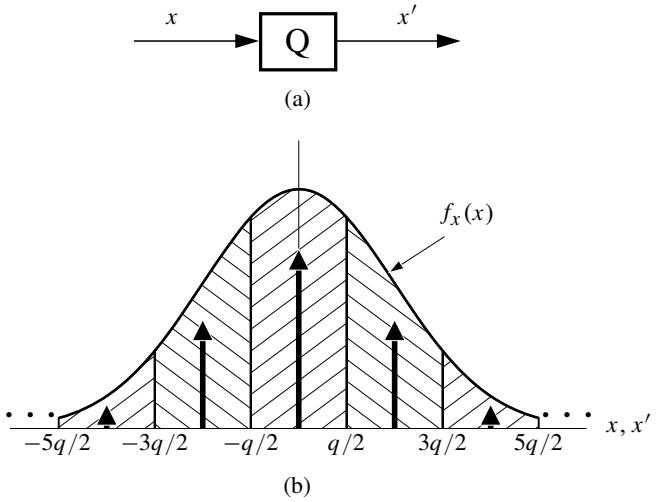

Figure 4.1 Formation of the PDF of the quantizer output x': (a) the quantizer; (b) area sampling.

are labeled as $\ldots -3q/2$ to $-q/2$, $-q/2$ to $q/2$, $q/2$ to $3q/2$, $\ldots$. The area under $f_x(x)$ within each quantization box is compressed into a Dirac delta function. The set of delta functions comprises $f_{x'}(x')$, the PDF of the quantizer output, also shown in Fig. 4.1(b). The formation of $f_{x'}(x)$ from $f_x(x)$ is the sampling process. We call this "area sampling."

Area sampling differs from conventional sampling in an important way. Area sampling can be represented as (a) convolution with a rectangular pulse function, and (b) conventional sampling.

To see how this works, refer once again to Fig. 4.1(b). We can express the area samples of $f_x(x)$ as

$$f_{x'}(x) = \cdots + \delta(x+q) \int_{-\frac{3q}{2}}^{-\frac{q}{2}} f_x(x)\,\mathrm{d}x + \delta(x) \int_{-\frac{q}{2}}^{\frac{q}{2}} f_x(x)\,\mathrm{d}x + \delta(x-q) \int_{\frac{q}{2}}^{\frac{3q}{2}} f_x(x)\,\mathrm{d}x + \cdots$$

$$= \sum_{m=-\infty}^{\infty} \delta(x-mq) \int_{mq-\frac{q}{2}}^{mq+\frac{q}{2}} f_x(x)\,\mathrm{d}x\,. \tag{4.2}$$

Now define a rectangular pulse function as

$$f_n(x) = \begin{cases} \dfrac{1}{q}, & -q/2 < x < q/2 \\ 0, & \text{elsewhere.} \end{cases} \tag{4.3}$$

This function has a unit area. The convolution of this function with $f_x(x)$ is

$$f_n(x) \star f_x(x) = \int_{-\infty}^{\infty} f_n(x-\alpha) f_x(\alpha)\,\mathrm{d}\alpha = \int_{x-\frac{q}{2}}^{x+\frac{q}{2}} \frac{1}{q} \cdot f_x(\alpha)\,\mathrm{d}\alpha\,. \tag{4.4}$$

Next, we multiply the result of this convolution by the impulse train $c(x)$, defined as

$$c(x) \stackrel{\triangle}{=} \sum_{m=-\infty}^{\infty} q\delta(x-mq)\,. \tag{4.5}$$

The product is

$$\Big(f_n(x) \star f_x(x)\Big) \cdot c(x) = \sum_{m=-\infty}^{\infty} q\delta(x-mq) \int_{x-\frac{q}{2}}^{x+\frac{q}{2}} \frac{1}{q} f_x(\alpha)\,\mathrm{d}\alpha\,. \tag{4.6}$$

When multiplying delta functions with other functions, we note that

$$\delta(x-mq) \cdot g(x) = \delta(x-mq) \cdot g(mq)\,. \tag{4.7}$$

Accordingly, the product can be written as

$$\Big(f_n(x) \star f_x(x)\Big) \cdot c(x) = \sum_{m=-\infty}^{\infty} \delta(x-mq) \int_{mq-\frac{q}{2}}^{mq+\frac{q}{2}} f_x(\alpha)\,\mathrm{d}\alpha\,. \tag{4.8}$$

From Eqs. (4.8) and (4.2) we conclude that

$$f_{x'}(x) = \Big(f_n(x) \star f_x(x)\Big) \cdot c(x)\,, \tag{4.9}$$

and so we have demonstrated that area sampling is first convolution with a rectangular pulse function, then conventional sampling.

The area sampling process can be visualized with the help of Fig. 4.2. The PDF of x is shown in Fig. 4.2(a). The rectangular pulse function is shown in Fig. 4.2(b). Their convolution is shown in Fig. 4.2(c). The sampling impulse train is shown in Fig. 4.2(d). The PDF of the quantizer output x' is shown in Fig. 4.2(e). It is the product of the impulse train and the convolution of the rectangular pulse function

with the PDF of x. We have used linear operations common in the field of digital signal processing to relate the probability density function of x and x'.

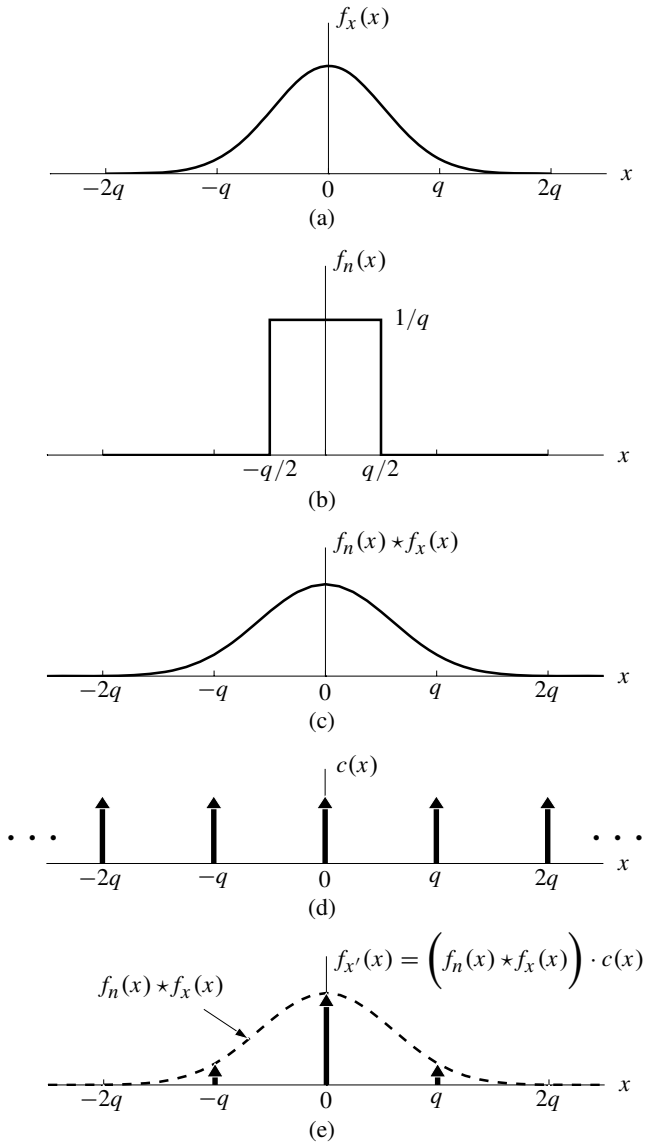

Figure 4.2 Derivation of PDF of x' from area sampling of the PDF of x: (a) PDF of x; (b) rectangular pulse function; (c) convolution of (a) and (b); (d) the impulse train; (e) PDF of x', the product of (c) and (d).

The Fourier transform of the rectangular pulse function is

$$\Phi_n(u) = \int_{-\infty}^{\infty} f_n(x) e^{jux}\, \mathrm{d}x = \int_{-q/2}^{q/2} \frac{1}{q} e^{jux}\, \mathrm{d}x = \operatorname{sinc} \frac{qu}{2}\,. \tag{4.10}$$

Working in the transform domain, we can use (4.10) to obtain the CF of x' from the CF of x.

The area sampling idea can be visualized in the CF domain by referring to Fig. 4.3. The CF of x, $\Phi_x(u)$, is shown in Fig. 4.3(a). The Fourier transform of the rectangular pulse function, $\operatorname{sinc}(qu/2)$, is shown in Fig. 4.3(b). The product of this transform and the CF of x, $\Phi_x(u) \cdot \operatorname{sinc}(qu/2)$, is shown in Fig. 4.3(c). (Multiplication in the CF domain corresponds to convolution in the PDF domain, as illustrated in Fig. 4.2.) This product is repeated in Fig. 4.3(d), and summed in Fig. 4.3(e) (repetition and summation in the CF domain corresponds to sampling in the PDF domain, as illustrated in Fig. 4.2).

The CF of x' is shown in Fig. 4.3(e). The repeated summed product can be represented by

$$\begin{aligned}\Phi_{x'}(u) &= \left(\Phi_x(u)\ \operatorname{sinc}\left(\frac{qu}{2}\right)\right) \star \left(\sum_{l=-\infty}^{\infty} \delta(u + l\Psi)\right) \\ &= \sum_{l=-\infty}^{\infty} \Phi_x\left(u + l\Psi\right) \operatorname{sinc}\left(\frac{q(u + l\Psi)}{2}\right),\end{aligned} \tag{4.11}$$

with

$$\Psi \triangleq \frac{2\pi}{q}\,. \tag{4.12}$$

Eq. (4.11) is an infinite sum of replicas. It is a periodic function of u, as would be expected for the Fourier transform of a string of uniformly spaced impulses. The replica centered about $u = 0$ is

$$\Phi_x(u) \cdot \operatorname{sinc} \frac{qu}{2}\,. \tag{4.13}$$

The derivation of Eq. (4.11) turns out to be a major result in quantization theory. In accord with this theory, the variable Ψ can be thought of as the "quantization radian frequency", expressed in radians per unit amplitude of x and x'. This is analogous to the sampling radian frequency Ω, expressed in radians per second. The "quantization period" is q, and this is analogous to the sampling period.

We can find a complete set of correspondences between the relevant quantities involved in sampling and quantization, respectively, by comparing Eqs. (4.11) and (4.5) to Eqs. (2.9) and (2.2). The correspondences are listed in Table 4.1.

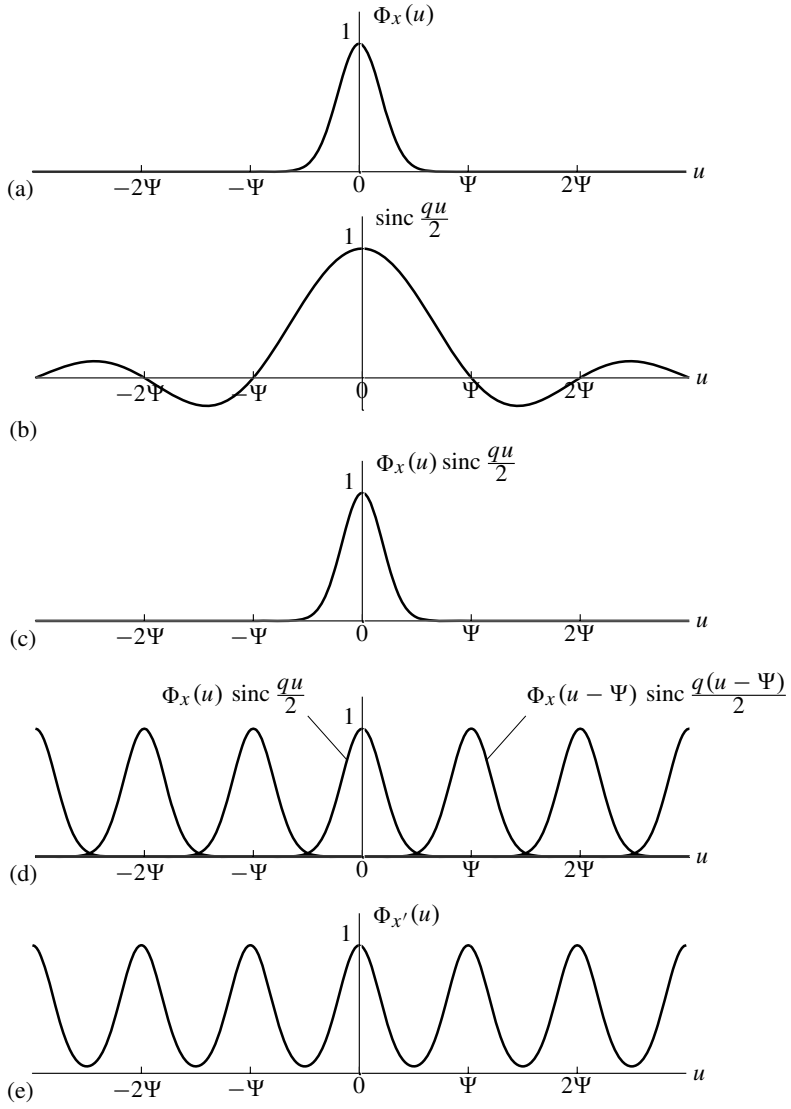

Figure 4.3 Formulation of area sampling in the CF domain: (a) CF of x; (b) CF of n, the sinc function; (c) CF of $x + n$; (d) the repetition of (c); (e) CF of x', the sum of the replicas.

4.2 COMPARISON OF QUANTIZATION WITH THE ADDITION OF INDEPENDENT UNIFORMLY DISTRIBUTED NOISE, THE PQN MODEL

At this time, it is useful to define an independent noise n, uniformly distributed between $\pm q/2$. This noise has zero mean and a variance of $q^2/12$. Its PDF is $f_n(n)$,

TABLE 4.1 Correspondences between sampling and quantization.

Eqs. (2.9) & (2.2)	k	t	n	ω	T	$\Omega = 2\pi/T$	Sampling
	$\Updownarrow$	$\Updownarrow$	$\Updownarrow$	$\Updownarrow$	$\Updownarrow$	$\Updownarrow$	$\Updownarrow$
Eqs. (4.11) & (4.5)	m	x	l	u	q	$\Psi = 2\pi/q$	Quantizing

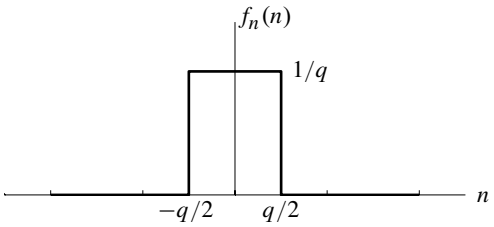

Figure 4.4 The PDF of an independent noise.

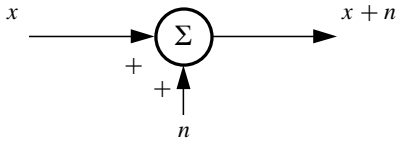

Figure 4.5 Addition of independent noise n to the signal x.

shown in Fig. 4.4. Let this noise be added to the signal x, as in Fig. 4.5. The result is $x + n$, whose PDF is the following convolution,

$$f_{x+n}(x) = f_n(x) \star f_x(x)\,. \tag{4.14}$$

Comparing Eqs. (4.14) and (4.9), we conclude that there is a fundamental relation between the PDF of the quantizer output and the PDF of the sum of x and the independent noise n. *The quantized PDF consists of samples of the PDF of x plus n.*

$$f_{x'}(x) = f_{x+n}(x) \cdot c(x)\,. \tag{4.15}$$

It is clear that quantization is not the same thing as addition of an independent noise. Fig. 4.6 shows two diagrams, one is quantization and the other is addition of independent uniformly distributed noise. The PDF of x' consists of uniformly spaced

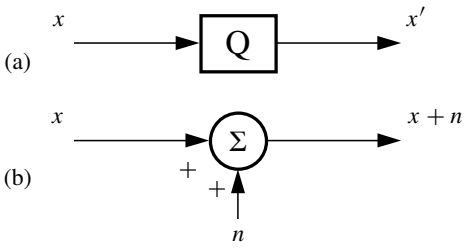

Figure 4.6 Comparison of quantization with the addition of uniformly distributed independent noises: (a) quantizer; (b) the "PQN model" of quantization.

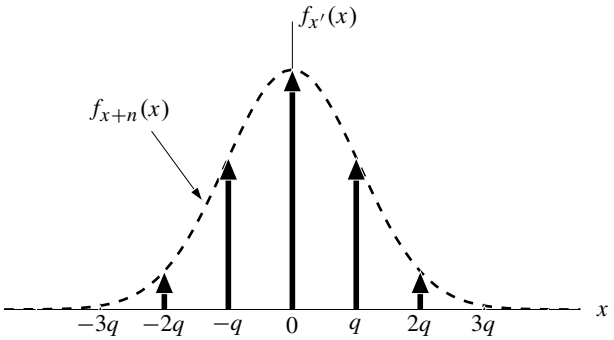

Figure 4.7 The PDF of x' consists of Dirac impulses, multiplied by q, and weighted by the samples of the PDF of $x + n$, exactly, under all conditions.

impulses. It is very different from the PDF of $x + n$, which is generally continuous and smooth.

To compare the PDF of x' with that of $x + n$, refer to Fig. 4.7. The PDF of x' is a string of Dirac impulses, corresponding to the samples of the PDF of $x + n$. This is true under all circumstances.

The CF of x', given by Eq. (4.11), is periodic with frequency $\Psi = 2\pi/q$. The CF of $x + n$ is aperiodic and is given by

$$\begin{aligned}\Phi_{x+n}(u) &= \Phi_x(u) \cdot \Phi_n(u) \\ &= \Phi_x(u) \cdot \operatorname{sinc} \frac{qu}{2} \,. \end{aligned} \tag{4.16}$$

This CF is replicated at frequency Ψ to form the CF of x' (refer to relations (4.11), (4.12), and (4.13)).

The noise n is an artificial noise. It is not the same as quantization noise. In subsequent chapters however, its statistical properties will be related to those of

quantization noise. It is natural therefore to call the noise n "pseudo quantization noise," or "PQN." Figure 4.6(b) depicts the "PQN model" that can be used for the analysis of quantization noise when appropriate quantizing theorems are satisfied.

4.3 QUANTIZING THEOREMS I AND II

If the replicas contained in $\Phi_{x'}(u)$ do not overlap, then $\Phi_x(u)$ can be recovered from $\Phi_{x'}(u)$. The Widrow quantizing theorem (Widrow, 1956a), analogous to the Nyquist sampling theorem, can be stated as follows:

Quantizing Theorem I (QT I)
If the CF of x is "bandlimited," so that

$$\Phi_x(u) = 0, \qquad |u| > \frac{\pi}{q} = \frac{\Psi}{2}, \tag{4.17}$$

then:

- *the replicas contained in $\Phi_{x'}(u)$ will not overlap*
- *the CF of x can be derived from the CF of x'*
- *the PDF of x can be derived from the PDF of x'.*

For the conditions of Quantizing Theorem I to be satisfied, the frequency of quantization must be more than twice as high as the highest "frequency component"[2] of $\Phi_x(u)$. The quantization grain size q must be made small enough for this to happen (making q smaller raises the "quantization frequency," spreads the replicas, and tends to reduce their overlap). How to obtain $f_x(x)$ from $f_{x'}(x')$ when QT I is satisfied will be described below.

If the replicas contained in $\Phi_{x'}(u)$ overlap but not enough to affect the derivatives of $\Phi_{x'}(u)$ at the origin in the CF domain, i.e. at $u = 0$, then the moments of x are recoverable from the moments of x'. Therefore, another quantizing theorem (Widrow, 1961) can be stated as follows:

Quantizing Theorem II (QT II)
If the CF of x is bandlimited so that

[2]Precisely, no overlap requires that $\Phi_x(u)$ is zero also at $|u| = \pi/q$. However, for characteristic functions, the only way that nonzero values at discrete points can make any difference in the PDF is that there are Dirac delta functions at these points. But no CF may contain Dirac delta functions, because this would mean that the integral of the PDF is infinite. Therefore, it is enough to prescribe that $\Phi_x(u)$ equals zero for $|u| > \pi/q$.

$$\Phi_x(u) = 0 \quad for \quad |u| > \frac{2\pi}{q} = \Psi\,, \tag{4.18}$$

then the moments of x can be calculated from the moments of x'.

For the condition for Quantizing Theorem II to be satisfied, the frequency of quantization must be higher[3] than the highest frequency component of $\Phi_x(u)$. The quantization grain size q must be made small enough for this to happen. The grain size q could be up to twice as large as the largest q that would still satisfy QT I. How to obtain the moments of x from the moments of x' when QT II is satisfied will be described below.

It is worth mentioning that if the conditions for QT I or QT II are satisfied, adding a mean μ to x allows these QTs to still be satisfied. This can be deduced from the expression for the CF of $x + \mu$:

$$\Phi_{x+\mu}(u) = e^{ju\mu}\,\Phi_x(u)\,. \tag{4.19}$$

Similarly, an arbitrary shift of the quantization transfer characteristic will not affect the fulfillment of the conditions of the theorems, either, see Exercise 4.21.

Quantizing Theorem II is unique to quantization theory and is not analogous to the sampling theorem. It should be noted that if QT I is satisfied, then QT II is automatically satisfied.

Other quantizing theorems will be introduced below, Quantizing Theorems III and IV, and they too are unique and not analogous to the Nyquist sampling theorem.

4.4 RECOVERY OF THE PDF OF THE INPUT VARIABLE x FROM THE PDF OF THE OUTPUT VARIABLE x'

If the quantization granularity is fine enough to satisfy the conditions of Quantizing Theorem I, then the PDF of the unquantized signal x can be recovered from the PDF of the quantized signal x'. It is useful when one only has quantized data, but needs the statistical properties of the original data.

[3]In strict sense, the above argumentation has proved QT II for the case when the derivatives of the CF are not altered at $u = 0$ by the repetitions due to quantization (see Eq. 4.11). This would require that all the derivatives of the CF equal zero for $|u| = \Psi$, that is, $\Phi(u)$ equals zero not only here, but also in an (arbitrarily small) environment of $|u| = \Psi$. This is not very important in the practical sense, but it is interesting from the theoretical point of view. It is not clear however whether this is really necessary. We do not have a formal proof yet but we were not able to construct a counterexample. For the examples enumerated in Appendix A, the corresponding derivatives at Ψ are zero for all the *existing* moments of x.

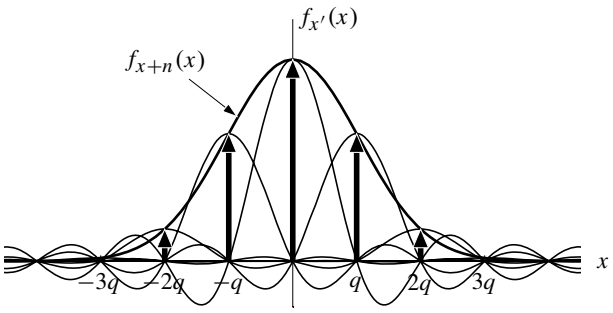

Figure 4.8 Sinc interpolation of $f_{x'}(x)$ to obtain $f_{x+n}(x)$ when Quantizing Theorem I is satisfied.

From the PDF of x', one can obtain the PDF of the sum of x and n, $f_{x+n}(x)$, by sinc function interpolation, illustrated in Fig. 4.8. The PDF of x can then be obtained from the PDF of $x + n$ by deconvolving $f_{x+n}(x)$ with $f_n(x)$. The deconvolution can be done in the PDF domain or in the CF domain. If done in the CF domain, one takes the CF of $f_{x+n}(x)$ and divides by the Fourier transform of $f_n(x)$, a sinc function, and the quotient is Fourier transformed to obtain $f_x(x)$. In the PDF domain, there are different ways of interpolation and deconvolution. A heuristic derivation of one such method will be described next (Widrow, 1956a).

We begin with a graphical method for obtaining $f_x(x)$ from $f_{x'}(x)$. Refer to Fig. 4.9. In Fig. 4.9(a), there is a sketch of $f_x(x)$. The running integral of $f_x(x)$ from $-\infty$ is defined as the cumulative distribution function,

$$F_x(x) = \int_{-\infty}^{x} f_x(\alpha)\,\mathrm{d}\alpha\,. \tag{4.20}$$

This function is sketched in Fig. 4.9(b). Also drawn in this figure are $F_x(x + q/2)$, the cumulative distribution function "advanced" by $q/2$, and $F_x(x - q/2)$, the cumulative distribution function "delayed" by $q/2$. Fig. 4.9(c) shows a curve which is the difference between $F_x(x + q/2)$ and $F_x(x - q/2)$, multiplied by $1/q$. It turns out that the resulting function is

$$\frac{1}{q}\Big(F_x(x + q/2) - F_x(x - q/2)\Big) = f_x(x) \star f_n(x) = f_{x+n}(x)\,. \tag{4.21}$$

Sampling $f_{x+n}(x)$ in Fig. 4.9(d) yields the PDF of the quantizer output, $f_{x'}(x)$. The discrete PDF $f_{x'}(x)$ is composed of area samples of $f_x(x)$.

Figure 4.10 is a "signal flow" block diagram that shows how $f_{x'}(x)$ is formed from $f_x(x)$. The steps involved correspond to the processing illustrated in Fig. 4.9.

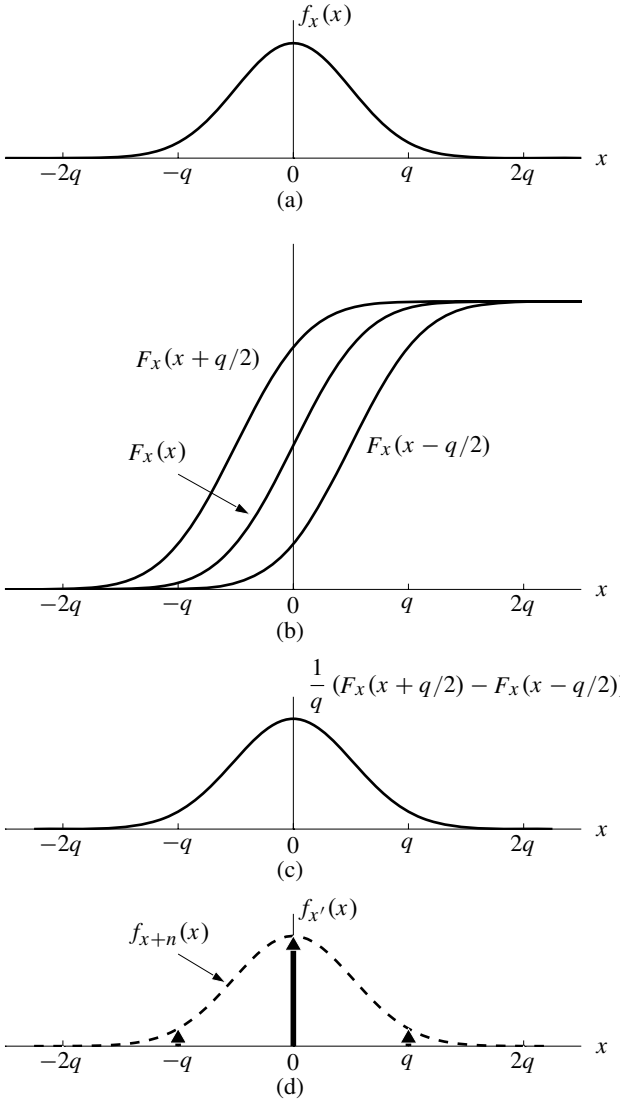

Figure 4.9 Relations between $f_x(x)$, $f_{x+n}(x)$, and $f_{x'}(x)$: (a) $f_x(x)$; (b) the cumulative distribution function $F_x(x)$ (the CDF), and "advanced" and "delayed" versions of it; (c) the scaled difference between the advanced and delayed cumulative distribution functions, $1/q \cdot (F(x+q/2) - F(x-q/2)) = f_{x+n}(x)$; (d) $f_{x'}(x)$, the samples of $f_{x+n}(x)$.

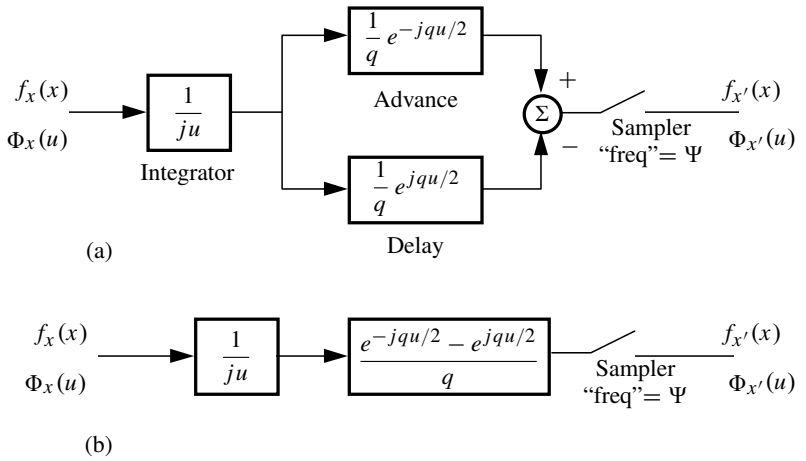

Figure 4.10 Getting $f_{x'}(x)$ from $f_x(x)$: (a) "signal flow" block diagram; (b) simplified block diagram.

First there is integration, then an advance operator and a delay operator[4] (with a scaling of $1/q$), and the difference is sampled to obtain $f_{x'}(x)$. Its transform is of course $\Phi_{x'}(u)$.

To recover $f_x(x)$ from $f_{x'}(x)$ when QT I is satisfied, the process can be reversed. Fig. 4.11 is a block diagram of a reversal process that works when conditions for QT I are met. The details are shown in Fig. 4.11(a). The sampling process of Fig. 4.10 is reversed by ideal lowpass filtering using the "sinc" interpolator. The advance and delay operators and their difference are reversed by their reciprocal transfer function in Fig. 4.11(a). The integrator in Fig. 4.10 is reversed by the differentiator in Fig. 4.11(a). Fig. 4.11(b) shows another flow diagram that is algebraically equivalent to the flow diagram of Fig. 4.11(a). Here we replace the reciprocal transfer function with its equivalent, a delay of $q/2$ and a discrete integrator. Fig. 4.11(c) is a flow diagram containing the same transfer function elements, but arranged sequentially to correspond to the actual operations to be done in the PDF domain.[5]

Figure 4.12 demonstrates the implementation of the process of recovering $f_x(x)$ from $f_{x'}(x)$. This process is based on the flow diagram of Fig. 4.11(c). The case illustrated in Fig. 4.12 is that of a Gaussian signal which has been very roughly

[4]Since the CF is defined as in (2.17), the advance and delay operators are $e^{-jqu/2}$, and $e^{jqu/2}$, respectively.

[5]Since the CDF is not absolutely integrable, these interchanges would require mathematical justification. Let us content ourselves here with the remark that in general, all of these operations are linear, and for absolutely integrable functions, they are also commutable. Heuristically, we would expect that commutability also holds for this case, since the CDF "behaves well," because it is the running integral of the PDF which is an absolutely integrable, nonnegative function.

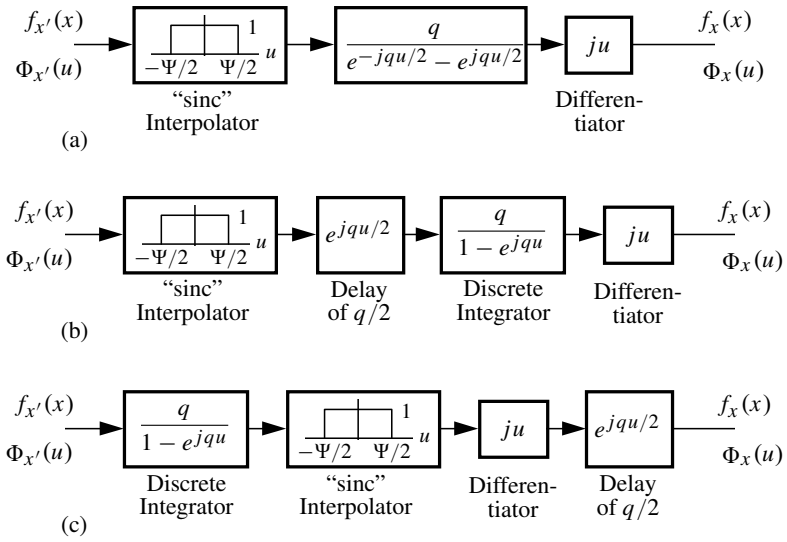

Figure 4.11 Getting $f_x(x)$ from $f_{x'}(x)$ when QT I is satisfied: (a) flow diagram; (b) equivalent flow diagram; (c) reordered flow diagram.

quantized, to a granularity of $q = 2\sigma$. Since 99.7 percent of the area of the PDF is contained between $\pm 3\sigma$, the histogram contains essentially three "bars" and the quantized PDF has essentially three impulses. The following steps have been taken:

- create new impulses whose values are a running sum of the values of the Dirac deltas of $f_{x'}(x)$ shown in Fig. 4.12(a),
- sinc interpolate the running sum impulses to get the cumulative distribution function, as shown in Fig. 4.12(b),
- differentiate the cumulative distribution function in order to obtain the PDF,
- shift left by $q/2$ in order to get the correct bias for the PDF, done in Fig. 4.12(c).

This method may seem complicated but it is in fact very straightforward. Proofs of the method under somewhat more restrictive conditions than those of QT I are given in the Addendum (available at the book's website), along with some numerical considerations.

A comparison of the true PDF in Fig. 4.12(c) with that interpolated from $f_{x'}(x)$ shows some error. The reason for the error is that the histogram is not perfectly represented with only three bars, and that the Gaussian CF is not perfectly bandlimited, but is very close to bandlimited for all practical purposes. In spite of this, the error is remarkably small.

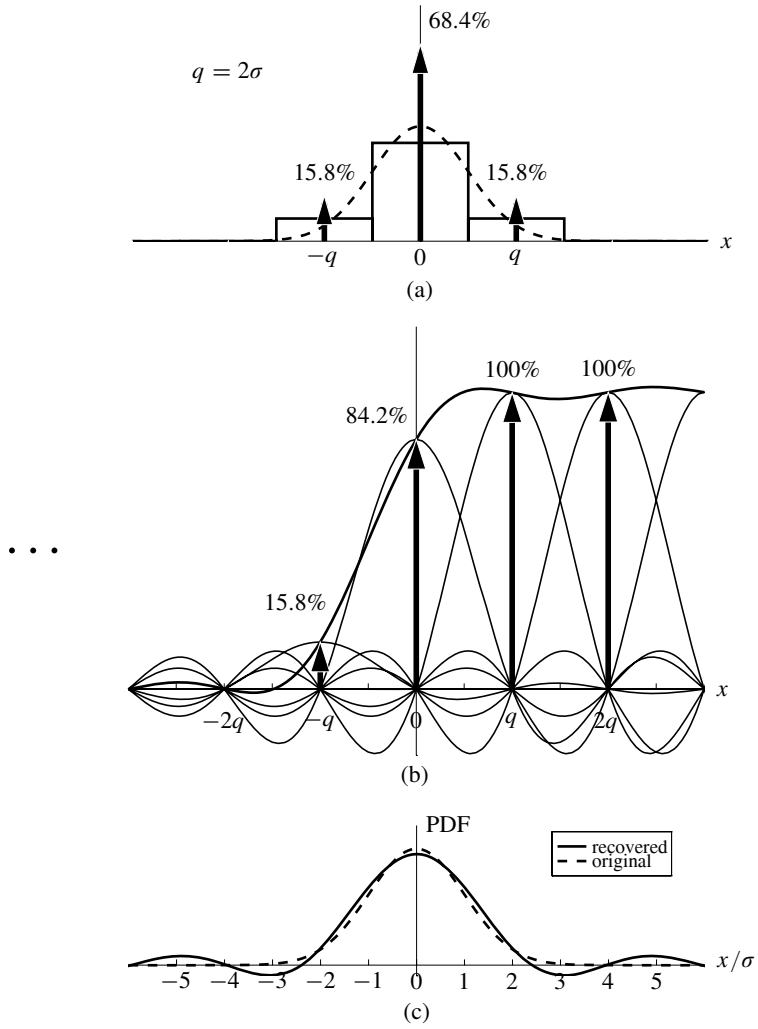

Figure 4.12 Getting the original PDF from the quantized PDF: (a) crude histogram of a Gaussian process with $q = 2\sigma$; (b) summing and sinc interpolation; (c) derivative of (b), shifted by $q/2$.

.

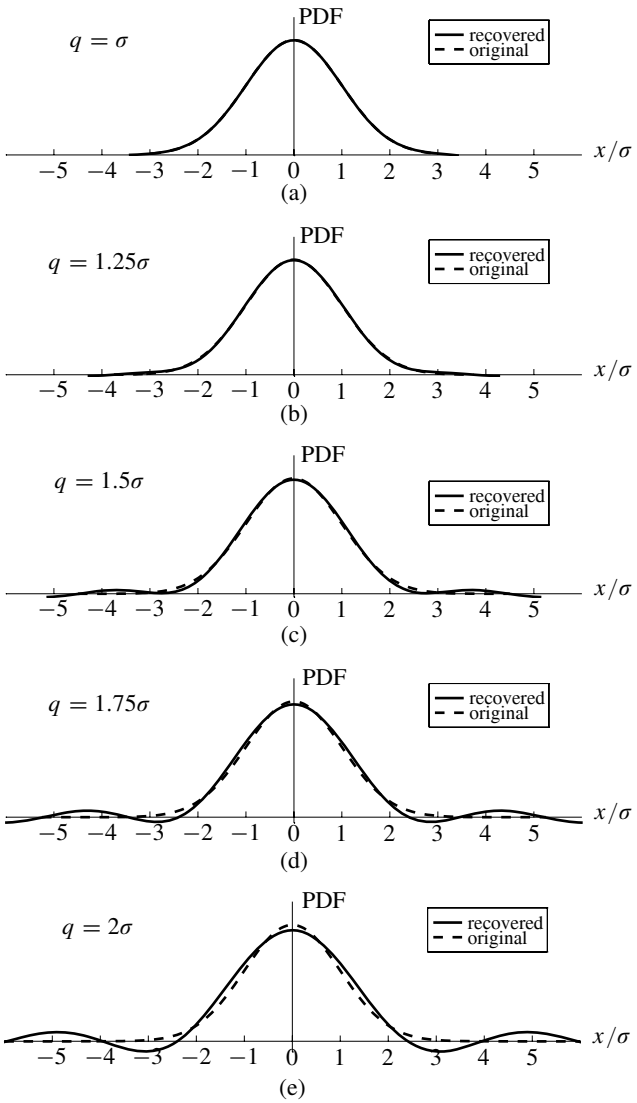

Figure 4.13 Reconstruction of the Gaussian PDF from a crude histogram, for different values of q.

Example 4.1 Reconstruction of PDF from Finely Quantized Data
Similar reconstructions have been done with finer quantization. The results are shown in Fig. 4.13. With $q = \sigma$, the errors are imperceptible. With $q = 1.25\sigma$, only the slightest error can be seen. With $q = 1.5\sigma$, a small error can be seen. The error is noticeable for $q = 1.75\sigma$, and for $q = 2\sigma$, the error is more noticeable, as we have already seen in Fig. 4.12(c).

Very precise interpolation only requires $q \leq \sigma$. This is a surprising result. When $q = \sigma$, the quantization is still very coarse.

Example 4.2 Interpolation of USA Census Data
Another histogram interpolation was undertaken using the USA 1992 census data. This was done as an exercise to test the effectiveness and limitations of the reconstruction process described above.

The 1990 USA census counted all citizens by age with 1-year intervals up to age 84. From there, people were counted with 5-year intervals up to age 99. All people of age 100 and greater were counted as a single group. The data that was available to us (Statistical Abstract, 1994) was statistically updated from the 1990 census to be applicable to the year 1992.

From the 1992 data, we constructed three crude histograms, the first with 5-year intervals, the second with 10-year intervals, and the third with 20-year intervals. Constructing these histograms amounted to quantizing the age variable with values of $q = 5$ years, $q = 10$ years, and $q = 20$ years. The question is, if the data were taken originally with ages roughly quantized, could a smooth population distribution be recovered?

Fig. 4.14(a) shows a 10-year interval histogram, and the interpolated distribution. Because the true population density has a discontinuity at age zero, it is not at all bandlimited. The interpolation process is therefore not really valid. The interpolated distribution function exhibits Gibbs' phenomenon at the discontinuity. The ringing corresponds to, among other things, negative and positive populations of people with negative ages.

The discontinuity can be eliminated by interpolating the histogram together with its mirror image. The result is shown in Fig. 4.14(b). The portion of the interpolated function with negative ages is to be disregarded.

Going back to the original interpolation, Fig. 4.14(c) shows a comparison of its right-hand part with the 1-year interval 1992 data. Gibbs' phenomenon prevents the interpolation from closely fitting the data. But comparing the right-hand part of the interpolation of the histogram and its mirror image with the 1-year 1992 data, the fit is seen to be very close. So, the census data could have been taken with 10-year age grouping, and a fine interpolation could have been made from it.

Interpolating histograms and their mirror images, tests were done to empirically determine the effects of roughness of quantization. Figs. 4.15(a),(b),(c) are histograms with 5-year, 10-year, and 20-year intervals. These histograms

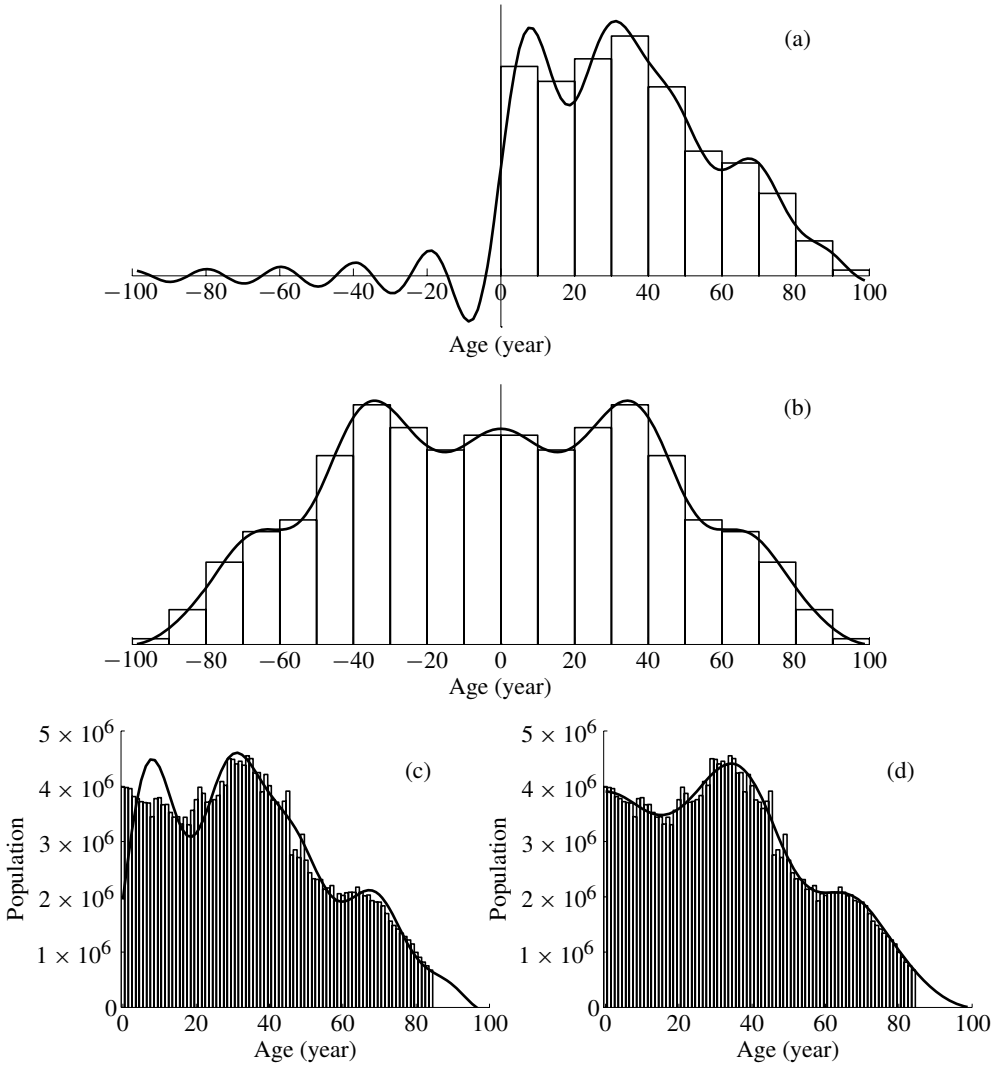

Figure 4.14 Reconstruction of USA age distribution from histogram with 10-year intervals, with the 1992 census data: (a) interpolation of histogram; (b) interpolation of histogram and its mirror image; (c) comparison of interpolation of histogram with one-year interval census data; (d) comparison of interpolation of histogram and its mirror image with 1-year interval census data.

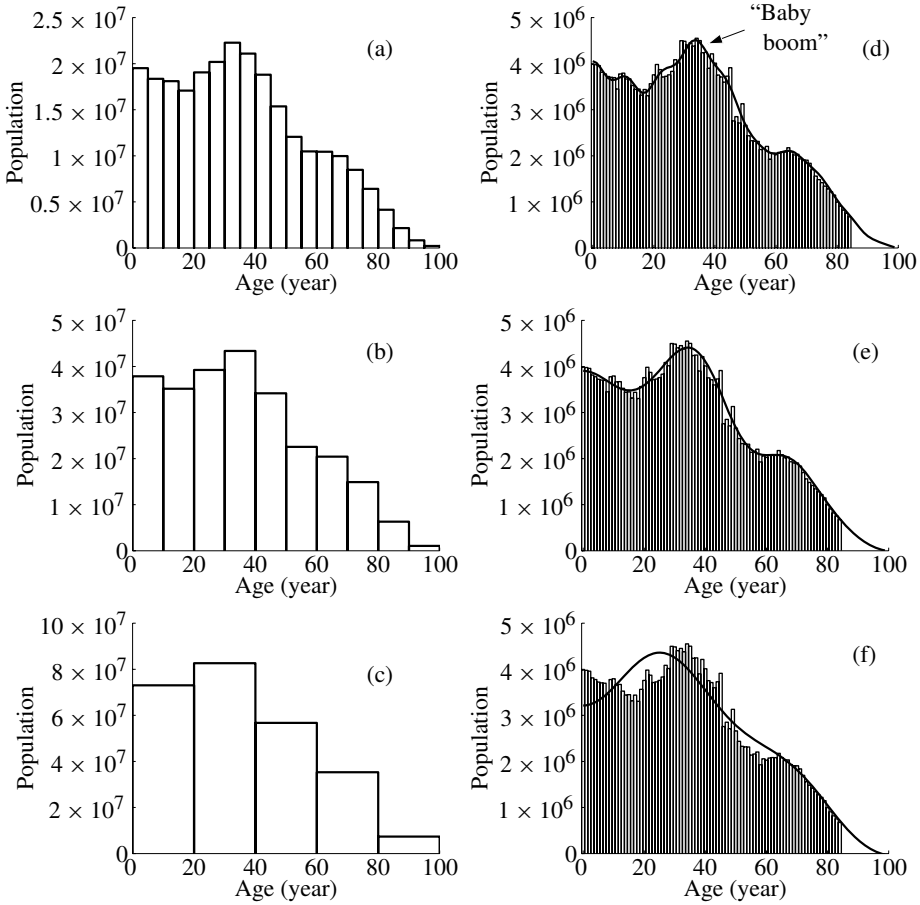

Figure 4.15 Reconstruction of the USA age distribution from histograms based on the 1992 census data: (a) 5-year interval histogram; (b) 10-year histogram; (c) 20-year histogram; (d) interpolation of the 5-year histogram; (e) interpolation of the 10-year histogram; (f) interpolation of the 20-year histogram.

and their mirror images were interpolated, and the right-hand portions are plotted in Figs. 4.15d, e, and f. With 5-year intervals, the interpolation fits the actual population data very closely. With 10-year intervals, the interpolation is more smoothed, but it still fits the actual data very closely. With 20-year intervals, the data is so coarse that the interpolation does not fit the actual data very well.

Demographers have for a long time been talking about the "baby boom." This is an anomalous bump in the USA population distribution from about age 25 to age 46 that shows up in the 1992 data. The big discontinuity at age 46 corresponds to the end of World War II in 1945, when American soldiers returned home and

from then on "made love, not war." The baby boom is caught very nicely in the interpolated function even when age is quantized to 10-year intervals.

4.5 RECOVERY OF MOMENTS OF THE INPUT VARIABLE x FROM MOMENTS OF THE OUTPUT VARIABLE x' WHEN QT II IS SATISFIED; SHEPPARD'S CORRECTIONS AND THE PQN MODEL

If the quantization granularity is fine enough to satisfy the conditions of Quantizing Theorem II, then the moments of x can be obtained from the moments of x'. Quantization theory can be used in a simple way to relate these moments.

The CF of x' contains an infinite number of replicas, and is given by Eq. (4.11) as

$$\Phi_{x'}(u) = \sum_{l=-\infty}^{\infty} \Phi_x\,(u + l\Psi)\,\mathrm{sinc}\left(\frac{q(u + l\Psi)}{2}\right). \tag{4.11}$$

This characteristic function is sketched in Fig. 4.3(d). If the conditions of Quantizing Theorem II are satisfied, then the replicas may or may not overlap, but if they do, there would be no overlap extending all the way to $u = 0$. The derivatives of the characteristic function at $u = 0$ would not be affected by overlap. Since the individual replica of (4.11) centered at $u = 0$ is given by Eq. (4.13),

$$\Phi_x(u) \cdot \mathrm{sinc}\,\frac{qu}{2}, \tag{4.13}$$

the derivatives of (4.13) at $u = 0$ would be the same as those of (4.11) when Quantizing Theorem II is satisfied. The moments of x' can, under these conditions, be obtained from the derivatives of (4.13) at $u = 0$.

Refer back to Fig. 4.6, comparing x' with $x + n$. The question is, how do their moments compare? The CF of $x + n$ is given by Eq. (4.16) as

$$\Phi_{x+n}(u) = \Phi_x(u) \cdot \mathrm{sinc}\,\frac{qu}{2}. \tag{4.16}$$

This is identical to (4.13). Therefore, when conditions for satisfying Quantizing Theorem II are met, the derivatives of (4.13) at $u = 0$ are identical to the corresponding derivatives of (4.16) and therefore the corresponding moments of x' are identical to those of $x + n$.

The variables x' and $x + n$ are quite different. The PDF of $x + n$ is continuous, while the PDF of x' is discrete. Under all circumstances, the PDF of x' consists of the Dirac samples of the PDF of $x + n$. Under special circumstances, i.e. when conditions for Quantizing Theorem II are met, the moments of x' are identical to

the corresponding moments of $x + n$. Then the key to the recovery of the moments of x from the moments of x' is the relationship between the moments of x and the moments of $x+n$. This relationship is easy to deduce, because n is independent of x, and its statistical properties are known. The mean and all odd moments of n are zero. The mean square of n is $q^2/12$, and the mean fourth of n is $q^4/80$. These facts can be verified for the rectangular PDF of n by making use of the formulas in Chapter 3 that defined the moments.

When the conditions of QT II are met, the first moment of x' is

$$\begin{aligned} \mathrm{E}\{x'\} = \mathrm{E}\{(x+n)\} &= \mathrm{E}\{x\} + \mathrm{E}\{n\} \\ &= \mathrm{E}\{x\}\,. \end{aligned} \tag{4.22}$$

The mean of x' equals therefore the mean of x. The second moment of x' is

$$\begin{aligned} \mathrm{E}\{(x')^2\} = \mathrm{E}\{(x+n)^2\} &= \mathrm{E}\{x^2\} + 2\mathrm{E}\{xn\} + \mathrm{E}\{n^2\} \\ &= \mathrm{E}\{x^2\} + \frac{1}{12}q^2\,. \end{aligned} \tag{4.23}$$

The third moment of x' is

$$\begin{aligned} \mathrm{E}\{(x')^3\} = \mathrm{E}\{(x+n)^3\} &= \mathrm{E}\{x^3\} + 3\mathrm{E}\{x^2n\} + 3\mathrm{E}\{xn^2\} + \mathrm{E}\{n^3\} \\ &= \mathrm{E}\{x^3\} + \frac{1}{4}\mathrm{E}\{x\}q^2\,. \end{aligned} \tag{4.24}$$

The fourth moment of x' is

$$\begin{aligned} \mathrm{E}\{(x')^4\} = \mathrm{E}\{(x+n)^4\} &= \mathrm{E}\{x^4\} + 4\mathrm{E}\{x^3n\} + 6\mathrm{E}\{x^2n^2\} + 4\mathrm{E}\{xn^3\} + \mathrm{E}\{n^4\} \\ &= \mathrm{E}\{x^4\} + \frac{1}{2}\mathrm{E}\{x^2\}q^2 + \frac{1}{80}q^4\,. \end{aligned} \tag{4.25}$$

The fifth moment of x' is

$$\begin{aligned} \mathrm{E}\{(x')^5\} = \mathrm{E}\{(x+n)^5\} &= \mathrm{E}\{x^5\} + 5\mathrm{E}\{x^4n\} + 10\mathrm{E}\{x^3n^2\} \\ &\quad + 10\mathrm{E}\{x^2n^3\} + 5\mathrm{E}\{xn^4\} + \mathrm{E}\{n^5\} \\ &= \mathrm{E}\{x^5\} + \frac{5}{6}\mathrm{E}\{x^3\}q^2 + \frac{5}{80}\mathrm{E}\{x\}q^4\,. \end{aligned} \tag{4.26}$$

Higher moments can be obtained by continuing in like manner.

In obtaining the above equations, we used the knowledge that the expected value of the sum is the sum of the expected values, and that the expected value of the product is the product of the expected values, since x and n are statistically independent.

The above expressions have the common form

$$\mathrm{E}\{(x')^r\} = \mathrm{E}\{x^r\} + M_r\,, \tag{4.27}$$

where M_r is the *moment difference*. Equation (4.27) holds only when the PQN model applies perfectly, when the conditions for QT II are met. Equation (4.27) depends on q, and the lower moments of x for $1, 2, \ldots r-2$.

$$\begin{aligned} M_1 &= 0 \\ M_2 &= \frac{q^2}{12} \\ M_3 &= \frac{1}{4}\mathrm{E}\{x\}q^2 \\ M_4 &= \frac{1}{2}q^2\mathrm{E}\{x^2\} + \frac{1}{80}q^4 \\ M_5 &= \frac{5}{6}\mathrm{E}\{x^3\}q^2 + \frac{5}{80}\mathrm{E}\{x\}q^4 \\ &\vdots \end{aligned} \tag{4.28}$$

In practice, moments of x' can be directly estimated from the quantized data. We are often interested in moments of x, the moments of the unquantized data. By using the above relations, the moments of x may be expressed in terms of the moments of x' as follows. The general expression that holds when the PQN model applies perfectly is

$$\mathrm{E}\left\{x^r\right\} = \mathrm{E}\left\{(x')^r\right\} - S_r\,. \tag{4.29}$$

Accordingly,

$$\begin{aligned} \mathrm{E}\{x\} &= \mathrm{E}\{x'\} - (0) \\ \mathrm{E}\{x^2\} &= \mathrm{E}\{(x')^2\} - \left(\frac{1}{12}q^2\right) \\ \mathrm{E}\{x^3\} &= \mathrm{E}\{(x')^3\} - \left(\frac{1}{4}\mathrm{E}\{x'\}q^2\right) \\ \mathrm{E}\{x^4\} &= \mathrm{E}\{(x')^4\} - \left(\frac{1}{2}q^2\mathrm{E}\{(x')^2\} - \frac{7}{240}q^4\right) \\ \mathrm{E}\{x^5\} &= \mathrm{E}\{(x')^5\} - \left(\frac{5}{6}\mathrm{E}\{(x')^3\}q^2 - \frac{7}{48}\mathrm{E}\{x'\}q^4\right) \\ &\vdots \end{aligned} \tag{4.30}$$

The expressions in the parentheses are known as "Sheppard's corrections" to the moments. Sheppard's first, second, third, ... corrections may be written as

$$\begin{aligned} S_1 &= 0 \\ S_2 &= \frac{q^2}{12} \end{aligned}$$

$$
\begin{aligned}
S_3 &= \frac{1}{4}\mathrm{E}\{x'\}q^2 \\
S_4 &= \frac{1}{2}q^2\mathrm{E}\{(x')^2\} - \frac{7}{240}q^4 \\
S_5 &= \frac{5}{6}\mathrm{E}\{(x')^3\}q^2 - \frac{7}{48}\mathrm{E}\{x'\}q^4 \\
&\vdots
\end{aligned}
\tag{4.31}
$$

The moment differences M_1,M_2,... are exactly equal to the corresponding Sheppard corrections S_1,S_2,... when the PQN model applies. When this is not the case, the Ms and the Ss will differ.

Sheppard's corrections for grouping (quantization) are well known in statistics. We have derived Sheppard's corrections from a comparison between quantization and the addition of an independent uniformly distributed noise, assuming that the CF of x is bandlimited and meets the conditions of Quantizing Theorem II. Sheppard's derivation was based on different assumptions. His work did not involve the characteristic function. The concept of bandlimitedness had not been invented yet. He assumed that the PDF of x was smooth, meeting certain derivative requirements. Although the assumptions are different, the two theories lead to identical moment corrections.

Sheppard's famous paper was a remarkable piece of work, published in 1898. It is a rigorous and deep algebraic exercise, but it is not easy to follow, and does not give an easy understanding. His derivation depends upon the smoothness of the PDF of x as stated in (Sheppard, 1898, p. 355): *"In the cases which we have specially in view, the curve $z = f(x)$ touches the base $z = 0$, to a very high order of contact, at the extreme points $x = x_0$ and $x = x_p$. In such a case $f(x_0)$ and $f(x_p)$ and their first few differential coefficients are zero."* Our derivation of Sheppard's corrections also depends on the smoothness of the PDF, as manifested in the bandlimitedness of the CF. A comparison of the characteristic function method and Sheppard's approach is given in the Addendum, readable on the website of the book.[6].

Our derivation gives a perspective on quantization that is equally as rigorous as Sheppard's, and it gives insight for people who have had experience with digital signal processing and digital control systems. When conditions for QT II are satisfied, moments may be calculated by treating quantization as equivalent to the addition of an independent noise which is uniformly distributed between $\pm q/2$. This equivalence yields Sheppard's corrections. This equivalence was probably never known by Sheppard.

[6] http://www.mit.bme.hu/books/quantization/

4.6 GENERAL EXPRESSIONS OF THE MOMENTS OF THE QUANTIZER OUTPUT, AND OF THE ERRORS OF SHEPPARD'S CORRECTIONS: DEVIATIONS FROM THE PQN MODEL

When neither QT I nor QT II are satisfied perfectly, the moments of x' deviate from those predicted by the PQN model. In order to examine this, we need precise expressions for the moments of x' that are general and apply whether or not QT I or QT II are satisfied. This subject is treated in detail in Appendix B.

The results described in the next section are based on the definitions and equations from Appendix B.

4.7 SHEPPARD'S CORRECTIONS WITH A GAUSSIAN INPUT

For the zero-mean Gaussian case, Sheppard's corrections work remarkably well. Since the PDF of the input x is symmetric about zero, all of Sheppard's odd-numbered corrections are zero, and the errors in these corrections are zero. Errors in Sheppard's even-numbered corrections are not zero, but they only become significant when the quantization is very rough. Using Eqs. (B.11) and (B.12), the relative error in Sheppard's second correction S_2 has been computed for a variety of choices of quantization box size q. The results are given in Table 4.2. The error in S_2 becomes significant only when q is made as large as $q = 2\sigma$. Then, the ratio R_2/S_2 becomes 9.5%. Even this is quite small.

The relative errors in Sheppard's fourth correction were computed with Eq. (B.18), and they are presented in Table 4.3. The residual error in S_4 becomes significant only when q is as large as $q = 2\sigma$. Then the ratio R_4/S_4 has a value of 28%. Sheppard's first four corrections work accurately for values of q equal to 1.5σ or less, and work reasonably well even when $q = 2\sigma$. This is very rough quantization.

TABLE 4.2 Sheppard's second corrections and their residual errors for zero mean Gaussian input signals.

q	$\mathrm{E}\{x^2\}$	$\mathrm{E}\{(x')^2\} = \mathrm{E}\{x^2\}$	$+S_2$	$+R_2$	R_2/S_2
$q = 0.5\sigma$	σ^2	σ^2	$+\,0.021\sigma^2$	$-\,2.1 \cdot 10^{-34}\sigma^2$	$-9.9 \cdot 10^{-33}$
$q = \sigma$	σ^2	σ^2	$+\,0.083\sigma^2$	$-\,1.1 \cdot 10^{-8}\sigma^2$	$-1.3 \cdot 10^{-7}$
$q = 1.5\sigma$	σ^2	σ^2	$+\,0.19\sigma^2$	$-\,6.5 \cdot 10^{-4}\sigma^2$	$-3.5 \cdot 10^{-3}$
$q = 2\sigma$	σ^2	σ^2	$+\,0.33\sigma^2$	$-\,0.032\sigma^2$	-0.095
$q = 2.5\sigma$	σ^2	σ^2	$+\,0.52\sigma^2$	$-\,0.20\sigma^2$	-0.38
$q = 3\sigma$	σ^2	σ^2	$+\,0.75\sigma^2$	$-\,0.55\sigma^2$	-0.73

TABLE 4.3 Sheppard's fourth corrections and their residual errors for zero mean Gaussian input signals.

q	$\mathrm{E}\{x^4\}$	$\mathrm{E}\{(x')^4\} = \mathrm{E}\{x^4\}$	$+S_4$	$+R_4$	R_4/S_4
$q = 0.5\sigma$	$3\sigma^4$	$3\sigma^4$	$+ 0.13\sigma^4$	$+ 6.5 \cdot 10^{-32}\sigma^4$	$5.1 \cdot 10^{-31}$
$q = \sigma$	$3\sigma^4$	$3\sigma^4$	$+ 0.51\sigma^4$	$+ 8.5 \cdot 10^{-7}\sigma^4$	$1.7 \cdot 10^{-6}$
$q = 1.5\sigma$	$3\sigma^4$	$3\sigma^4$	$+ 1.2\sigma^4$	$+ 0.022\sigma^4$	0.019
$q = 2\sigma$	$3\sigma^4$	$3\sigma^4$	$+ 2.1\sigma^4$	$+ 0.59\sigma^4$	0.28
$q = 2.5\sigma$	$3\sigma^4$	$3\sigma^4$	$+ 3.0\sigma^4$	$+ 2.4\sigma^4$	0.79
$q = 3\sigma$	$3\sigma^4$	$3\sigma^4$	$+ 3.0\sigma^4$	$+ 4.8\sigma^4$	1.6

To summarize the findings of this section, we have determined for Gaussian inputs that the corrections needed to accurately utilize Sheppard's first four corrections are negligible, small, or only moderate for quantization grain sizes that could be as large as $q = 2\sigma$. The Gaussian inputs may be with or without mean values.

Sheppard's corrections work perfectly when QT II is satisfied. Sheppard's corrections work almost perfectly in the Gaussian case even with rough quantization, because the condition for QT II are approximately met. The effects of overlap of replicas in the CF domain are minimal because the Gaussian CF drops off with $e^{-|u^2|}$. Other CFs that do not drop off this rapidly correspond to cases where Sheppard's corrections are not so accurate when rough quantization is practiced.

4.8 SUMMARY

The purpose of this chapter has been to relate the probability density function of the quantizer output to that of its input, and vice versa. This chapter also relates the moments of the quantizer output to those of its input, and vice versa.

The PDF of the quantizer output is a string of uniformly spaced Dirac impulses. The spacing is the quantum step size q. This output PDF can be obtained by Nyquist sampling of the input PDF convolved with a uniform PDF which is distributed between $\pm q/2$.

When the quantum step size q is made small enough so that $2\pi/q$, the "quantization radian frequency," is at least twice as high as the highest "frequency" component contained in the Fourier transform of the input PDF (its characteristic function), the input PDF is perfectly recoverable from the output PDF. This is Widrow's quantizing theorem, now known as QT I, and it is based on Nyquist sampling theory applied to quantization.

The staircase input–output function that describes the uniform quantizer is a nonlinear function. Nevertheless, the quantizer output PDF is linearly related to the

input PDF. The quantizer operates nonlinearly on signals, but operates linearly on their probability densities.

When the quantization frequency is at least as high as the highest frequency component contained in the quantizer input, the moments of the quantizer input signal are perfectly recoverable from the moments of the quantizer output signal. This is Widrow's second quantizing theorem, now known as QT II, and it does not correspond to Nyquist's sampling theorem.

When the conditions for QT II are met, the moments of the quantizer output signal are identical to those of the sum of the quantizer input signal and an independent noise that is uniformly distributed between $\pm q/2$. This noise is called pseudo quantization noise, or PQN.

When one is only concerned with moments, it is often useful to compare quantization with the addition of independent PQN. The addition of PQN to the quantizer input signal is called the PQN model.

Although quantization is often represented in the literature in terms of the PQN model, these two processes are very different from one another. Quantization yields a PDF that is a string of uniformly spaced impulses, while addition of independent PQN to the quantizer input signal generally yields a smooth PDF (being a convolution of the input PDF with a PDF that is uniformly distributed between $\pm q/2$). Although the PDFs of the quantizer output and the output of the PQN model are very different, they have precisely corresponding moments when conditions for QT II are met.

When conditions for QT I are met, conditions for QT II are automatically met. When conditions for QT II are met, conditions for QT I are not necessarily met. Conditions for QT II are easier to meet than are conditions for QT I.

When conditions for QT II are met, the PQN model can be used to calculate the differences between the moments of a quantized variable and the corresponding moments of the unquantized variable. The moment differences are called Sheppard's corrections (Sheppard, 1898). They allow one to estimate the moments such as mean, mean square, etc. of a signal from its digitized samples. Sheppard derived these corrections long ago by making certain assumptions about the smoothness of the PDF of the variable before quantization. Our assumptions are quite different, based on bandlimitedness of the characteristic function. Both theories of quantization yield the same moment corrections when their respective conditions on the PDF are met.

Gaussian signals have Gaussian characteristic functions. This type of CF is clearly not bandlimited, so the PQN model is not perfect for moment calculations. Expressions for the errors in moment calculations have been derived in Appendix B for cases where conditions for QT II are not perfectly met. These expressions have been applied to the quantization of Gaussian signals. They reveal that moment predictions based on the PQN model are extremely accurate even when the quantization step size q is as big as σ, one standard deviation. Errors in Sheppard's mean square correction are of the order of one part in 10^7, and in the mean fourth correction are of

the order of two parts in 10^6. Thus, Sheppard's corrections and the PQN model work very well even for extremely rough quantization when one is working with Gaussian signals. For non-Gaussian signals, these models also work well with fine-to-rough quantization, as will be shown in Appendices G and H.

4.9 EXERCISES

4.1 Refer to Fig. 4.2 (page 64). Suppose that the PDF is Gaussian with $\sigma = 2q$, and generate the plot of Fig. 4.2(a). Using Matlab, convolve this with the function in Fig. 4.2(b) to get the rest of Fig. 4.2.

4.2 Suppose that the CF of Fig. 4.3(a) (page 66) is Gaussian. Using Matlab, calculate and plot the other functions shown in Fig. 4.3.

4.3 Let the original PDF in Fig. 4.12 (page 75) be Gaussian. Calculate the interpolated PDF in Fig. 4.12(c), and compare it with the original Gaussian PDF.

4.4 Calculate the functions shown in Fig. 4.8 (page 71) for a Gaussian input with zero mean.

4.5 Let the PDF $f(x)$ be rectangular in $(-0.8q, 0.8q)$. Derive analytically the functions in Fig. 4.9 (page 72). What effect does the mean and the width of the rectangular function have?

4.6 Do a Monte Carlo experiment to measure moments of the quantizer output. Compare the results with results obtained with the PQN model. Let the input signal be Gaussian with $\sigma = q$, and let the mean be $\mu_x = q/4$.

4.7 The 1990 census data of Hungary are available from the web page

`http://www.mit.bme.hu/books/quantization/exercises/.`

Follow the method of Section 4.4 as illustrated in Figs. 4.14–4.15 (pages 77–79) to recover the population distribution from a 10-year interval histogram (construct a 10-year interval histogram from the web data). Use the mirror image approach. Plot and compare the interpolated distribution with the original 1-year interval census data. Is the quality of reconstruction different from that obtained for the USA census data in Example 4.2?

4.8 Fig. E4.8.1 illustrates the quantization of a uniformly distributed random variable.

(**a**) Calculate and plot the PDF of the quantization noise. Calculate the mean and mean square of the quantization noise.

(**b**) From the histogram, use the interpolative method of Section 4.4 to calculate and plot the interpolated PDF. This assumes that the Nyquist condition was satisfied (of course, it was not satisfied). Plot the true PDF together with the interpolated PDF and compare them.

4.9 Fig. E4.9.1 illustrates the quantization of a triangularly distributed random variable.

Answer the same questions as in Exercise 4.8.

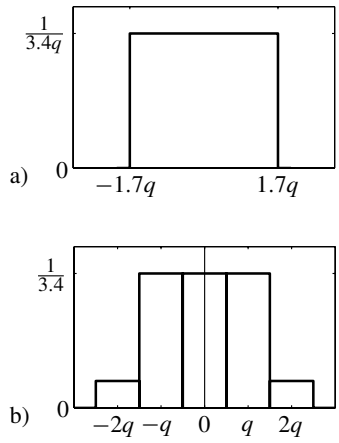

Figure E4.8.1 Quantization of a uniformly distributed random variable: (a) PDF of the input variable; (a) histogram of the quantized variable.

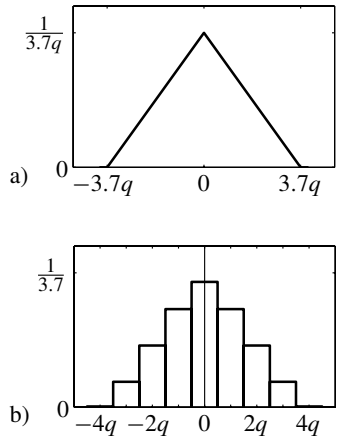

Figure E4.9.1 Quantization of a triangularly distributed random variable: (a) PDF of the input variable; (b) histogram of the quantized variable.

4.10 Given an input signal x with a symmetric triangular PDF in $\pm 2q$. Let this signal be quantized.

(**a**) Write an expression for the PDF and the CF of the quantized output signal.

(**b**) Derive the first four moments of x and x'.

(**c**) Apply Sheppard's corrections to x' to get the moments of x. How well do these work? How big are the errors in Sheppard's corrections compared to the corrections themselves? Find the ratios.

(**d**) Check the nonzero errors by Monte-Carlo.

4.11 Repeat the questions of Exercise 4.10 for x with uniform PDF in $(\pm A)$ with $A = (N + 0.5)q$, where N is a nonnegative integer.

4.12 Repeat the questions of Exercise 4.10 for x with uniform PDF in $(\pm A)$ with $A = Nq$, where N is a positive integer.

4.13 Repeat the questions of Exercise 4.10 for PDF of x equal to the "house" PDF (see Fig. E3.12.1, page 55), with $A = 2q$. Evaluate the errors numerically for for $\alpha = 1$.

4.14 Repeat the questions of Exercise 4.10 for PDF of x being the sum of two Gaussians:

$$f(x) = \frac{0.4}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{(x-\sigma)^2}{2\sigma^2}} + \frac{0.6}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{(x+\sigma)^2}{2\sigma^2}}\,, \tag{E4.14.1}$$

with $\sigma = q$.

4.15 Do the PDFs in Exercises 4.10, 4.11, 4.12, 4.13, 4.14 satisfy QT I? QT II?

4.16 For the random variable x, a series of samples are given in the file `samples_of_x.mat` available from the web page

`http://www.mit.bme.hu/books/quantization/exercises/`.

For $q = 1$, plot the PDF of x and x', slice up the PDF of x to get the PDF of v. Numerically calculate the first four moments of x and x'. Apply Sheppard's corrections to moments of x' to approximate moments of x. How well do Sheppard's corrections work here?

4.17 The random variable x' is the result of the quantization of x:

$$x' = m \qquad \text{if} \qquad m - 0.5 \le x < m + 0.5, \qquad m = 0, \pm 1, \ldots \tag{E4.17.1}$$

Here x is uniformly distributed, its mean value is $\mu = K + 0.25$ (K is an integer) the range of the probability density function is 3.

(a) Determine the distribution of x'.

(b) Give the mean and the variance of x'. How large is the increase in the variance due to the quantization?

4.18 Determine for the characteristic function defined in Exercise 3.13, which of Sheppard's corrections are exact (a) if $q = a$, or (b) if $q = a/2$.

4.19 For the characteristic function of the random variable x, the following equality is true:

$$\Phi_x(u) = 0 \qquad \text{if} \qquad |u| > \frac{2\pi}{q}\,. \tag{E4.19.1}$$

The uniformly quantized version of x is denoted by x'. Using the PQN model, express $\mathrm{E}\{x^6\}$ in terms of the moments of x' and q (sixth-order Sheppard correction).

4.20 A simple, perfectly bandlimited characteristic function is the triangle CF between $\pm B$, see Fig. E4.20.1:

$$\Phi(u) = \begin{cases} 1 - \dfrac{|u|}{B} & \text{if } |u| \le B, \\ 0 & \text{otherwise.} \end{cases} \tag{E4.20.1}$$

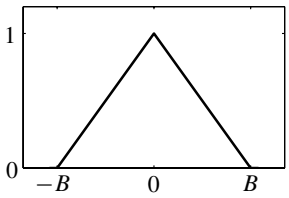

Figure E4.20.1 Triangle CF.

(**a**) What is the corresponding PDF?

(**b**) What can you say about the moments of the random variable?

The random variable is quantized with quantum size q. Plot the characteristic function of the quantized variable, and the PDF of the quantization error for

(**c**) $q = \pi/B$ (QT I is fulfilled),

(**d**) $q = 2\pi/B$ (QT II is fulfilled),

(**e**) $q = 3\pi/B$ (neither QT I nor QT II is fulfilled).

Can you determine the moments of the quantized variable?

4.21 The derivations in Chapter 4 analyze the mid-tread quantizer, with a "dead zone" around zero.

(**a**) Generalize the expression of the characteristic function of the output (4.11) to the case when the quantizer is shifted by s both horizontally and vertically (that is, the quantizer characteristic is shifted along the 45° line).

(**b**) Investigate the results for the case of the mid-riser quantizer ($s = q/2$).

(**c**) Show how the same result as in (a) can be obtained from (4.11) by one of the transformations illustrated in Fig. 1.2 (page 5).

(**d**) Are these cases different from the case of input offset? Why? Where does the output characteristic function differ?

(**e**) If you solved Exercise 2.10 (page 30), show how the result of this exercise can be directly applied to quantization.

4.22 An attractive way to obtain the $f_x(x)$ from $f_{x'}(x)$ is to do interpolation and deconvolution in the same step. Determine an interpolation formula between samples of $f_{x'}$ and $f_x(x)$.

4.23 Assume that for numbers equal to (integer + 0.5)q,

(**a**) either rounding towards zero (see page 12),

(**b**) or convergent rounding (see page 396) is implemented.

Describe the behavior of v for x uniformly distributed on the discrete values $\{(k + 0.5)q\}$, $k = 1, 2, ..., N$.

4.24 Derive the general formula of the rth moment of x' as a function of q and the moments of x. Hint: The methodology of Stuart and Ord (1994) is explained in the Addendum-readable in the website of the book.[7]

4.25 Derive the general expression of Sheppard's corrections, using the so-called Bernoulli numbers.[8]
Hint: Prove the following formula:

$$\mathrm{E}\{x^r\} = \mathrm{E}\{(x')^r\} - \left(\sum_{m=1}^{r} \binom{r}{m} \left(1 - 2^{1-m}\right) B_m q^m \mathrm{E}\{(x')^{r-m}\} \right) \qquad \text{(E4.25.1)}$$

(see (Stuart and Ord, 1994)).

4.26 Calculate the errors in Sheppard's corrections from expressions (B.8)–(B.16), and compare with the errors found in

(**a**) Exercise 4.10,
(**b**) Exercise 4.11,
(**c**) Exercise 4.12,
(**d**) Exercise 4.13.

4.27 Find the expression for the error in Sheppard's 5th correction (see Eqs. (B.3)–(B.6) and (B.8)–(B.16)).

4.28 For distributions selected from the ones described in Appendix I (A Few Properties of Selected Distributions), check numerically the asymptotic behavior of

(**a**) the arithmetic mean,
(**b**) the variance,
(**c**) the input to quantization noise correlation,
(**d**) the correlation coefficient between input and quantization noise.

Determine the forms and coefficients of the envelopes.

Hint: for numerical calculations, use the program `qmoments`, available from the web page of the book as part of the roundoff toolbox:

`http://www.mit.bme.hu/books/quantization/`.

[7] `http://www.mit.bme.hu/books/quantization/`

[8] The Bernoulli numbers (Stuart and Ord, 1994; Korn and Korn, 1968) are defined as the coefficients of $t^n/n!$ in the Taylor series of $t/(e^t - 1)$:

$$\frac{t}{e^t - 1} = \sum_{n=1}^{\infty} \frac{B_n t_n}{n!} \,. \qquad \textit{(4.25.FN1)}$$

$B_0 = 1$, $B_1 = -1/2$, $B_2 = 1/6$, $B_3 = 0$, $B_4 = -1/30$, $B_5 = 0$, $B_6 = 1/42$, and so on. A simple recurrence relationship allows the calculation of the Bernoulli numbers of higher index:

$$B_n = -\frac{1}{n+1} \sum_{k=0}^{n-1} \binom{n+1}{k} B_k \,, \qquad n = 1, 2, \ldots$$