DIGITAL SIGNAL PROCESSING

1. Digital signals

1.1. Introduction

Formally a digital signal is defined as a function of a discrete parameter, and of which the function values are taken from a finite set of values. It can therefore be denoted by $x(n)$ where n is a variable assuming integer values. Although such a definition is mathematically both correct and sufficient, the engineer mostly likes to have an interpretation of the term signal that gives him more feeling with practice. In most analogue signals the independent variable is denoted by t and is implicitly assumed to stand for time unless otherwise stated. Also for digital signals it is often customary to insert some notion of time into the formalism by denoting the signal values by $x(nT)$ where T is a quantity (with dimension seconds) that stands for sampling period, which is the time between the "occurrence" of two succeeding signal values (often then denoted by "samples"). Although the interpretation of a digital signal as a sequence of samples occurring at regular periods of time is quite convenient in practice, we will use throughout the notation $x(n)$, $y(n)$ etc. rather than $x(nT)$ or $y(nT)$, and keep track of the corresponding sampling period T by separately mentioning it where necessary.

Having elaborated on the first characteristics of a digital signal, i.e. its time-discrete nature, something must be said also about the finite set of possible function values. The use of digital signals has emerged from the possibilities of using digital computers for processing of signals. Such a computer operates on numbers, mostly encoded in a binary format, which means that every number is represented by a string of 0's and 1's. The length of such a string is fixed and finite. Nowadays, resulting from the advents of solid state technology, very small special purpose hardware devices (possibly even on a simple integrated circuit) have been developed to do the job. In these machines the length of the strings of 0's and 1's that represent the numbers is very limited. This length, often designated as wordlength, is typically 8, 12 or 16 bits, and therefore only 2^8, 2^{12} or 2^{16} different values can then be assumed by the number that is represented by such a string of 0's and 1's. This finite wordlength often is a complicating factor in the analysis of digital systems, and usually what one does is first neglecting this quantized character of the digital signals and analyzing the system assuming the signals to be discrete in time but not quantized in amplitude.

Most theory treated in textbooks is therefore devoted actually to discrete-time signals rather than to digital signals. Although this may give some confusion in the use of the terms digital signals and digital systems we will follow this convention, and actually discuss discrete time signals and systems, that is, treating digital signals as if they may have any amplitude value, and not bothering about their actual representation as "words" consisting of strings of 0's and 1's.

In a later stage of analysis one nevertheless has to cope with the finite wordlength of the signals in some way or another. Some of the effects caused by this finite wordlength are discussed in appendix A.

1.2. Representation in the time domain

A pictorial representation of a digital signal is given in fig. 1.1. Where convenient we will use a double indication on the "time"-axis, i.e. indicate the order number n and the corresponding sampling instant nT.

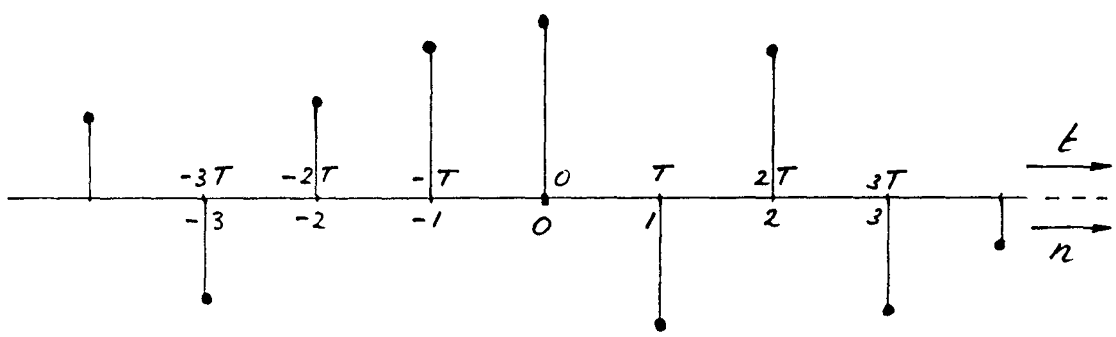

Fig. 1.1.

Unless otherwise stated it will always be assumed that the digital signal is defined for $n \varepsilon (-\infty, \infty)$. A digital signal is said to be of finite duration if for some integers N_1 and N_2.

$$x(n) = 0 \qquad \begin{cases} n < N_1 \\ n > N_2 \end{cases}$$

(Note that the signal is defined for all n.)

The unit impulse function $\delta(n)$ is defined by:

$$\delta(n) = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0 \end{cases} \tag{1.1}$$

It is depicted in fig. 1.2, and clearly is of finite duration.

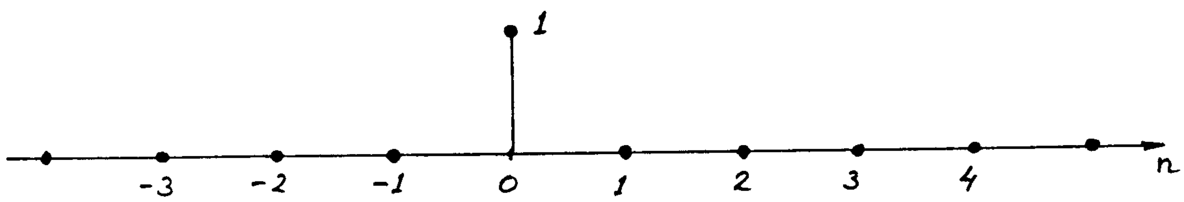

Fig. 1.2.

The unit step function of fig. 1.3 is defined by

$$u(n) = \begin{cases} 0 & n < 0 \\ 1 & n \geq 0 \end{cases} \qquad (1.2)$$

and is not of finite duration.

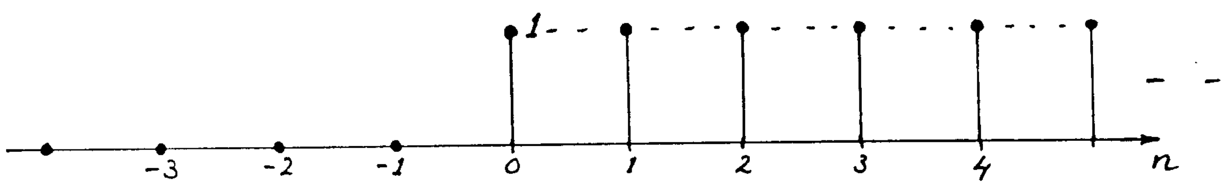

Fig. 1.3.

A discrete sinusoidal signal has the form

$$x(n) = A \sin (n\theta + \varphi) . \qquad (A > 0) \qquad (1.3)$$

A is the amplitude and φ an arbitrary phase. θ is the relative frequency and will play an important role in the analysis of digital systems.

A digital signal is said to be periodic of some period length N if N is the smallest integer for which

$$x(n+N) = x(n) \qquad \forall n. \qquad (1.4)$$

The first observation to be made is that if x(n) is periodic and not identically zero, then it cannot be of finite duration.

Now let us try to determine whether or not a sinusoidal signal is periodic. From (1.3) and (1.4) we have that

$$A \sin((n+N)\ \theta+\varphi) = A \sin (n\theta + \varphi)$$

which can only be satisfied for all n if

$$N\ \theta = k.2\pi$$

or

$$\theta = \frac{k}{N} \ .\ 2\pi \qquad (1.5)$$

where k is an arbitrary integer.

Therefore we can make the following observations

1) Not every sinusoidal digital signal is periodic.
   It is periodic only if its relative frequency is a rational fraction of 2π.
2) If the signal is periodic, and θ is thus given by (1.5) then the period may be extremely long even for not too small values of the relative frequency. Also a small change of the frequency may drastically change the period. As an example consider two signals with relative frequencies

   $$\theta_1 = \frac{1}{6} \cdot 2\pi, \; \theta_2 = \frac{7}{45} \cdot 2\pi$$

   respectively(see fig. 1.4). Their periods are 6 and 45 respectively.
3) If a sinusoidal signal is periodic, then all of its sample values may be less than the amplitude. Consider, for example the signal

   $$x(n) = 10 \cdot \sin (n \cdot \frac{1}{6} 2\pi)$$

   which is periodic with period length $\underline{N} = 6$ and has amplitude 10. It only assumes the values 0 and $\pm 5\sqrt{3}$ (see fig. 1.4).
   If, on the other hand, the signal is non-periodic, then it is always possible to find a sample $x(n)$ that has a value that is arbitrarily close to the amplitude A.

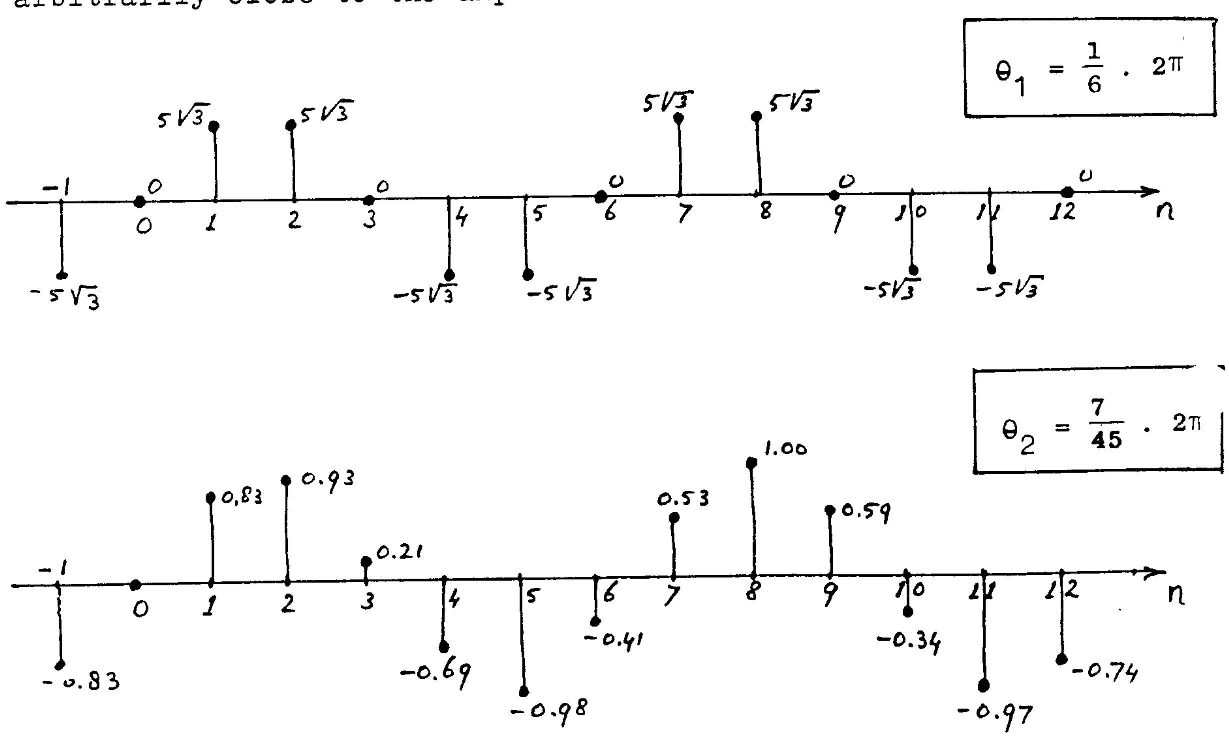

Fig. 1.4.

4) Consider two sinusoidal signals with relative frequency

   $$\theta_1 = \theta, \quad \theta_2 = 2\pi - \theta \quad , \text{ where } 0 < \theta < \pi$$

   and the same amplitude, thus

   $$x_1(n) = A.\sin n\, \theta_1 = A.\sin n\, \theta$$

   $$x_2(n) = A.\sin (n\, (2\pi - \theta)) = -A \sin n\, \theta$$

Except for the sign these two signals are equal from sample to sample:

$$x_1(n) = -x_2(n) \qquad \forall\, n.$$

Similarly consider a signal with rel.freq. $\theta_3 = 2\pi + \theta$

$$x_3(n) = A \sin (n\,(2\pi + \theta)) = A \sin n\,\theta$$

Also $x_3(n)$ is identical to $x_1(n)$. We see that signals with different relative frequency can nevertheless be identical and are therefore indistinguishable. In fact we have shown that with every θ outside the interval $[0,\pi]$ there corresponds a θ' inside this interval such that the corresponding sinusoids are indistinguishable. Therefore without loss of generality we may assume that the relative frequency is always in the interval

$$0 \leq \theta \leq \pi \tag{1.6}$$

A digital signal can be shifted in time (delayed or advanced) by changing the index number.

$x(n-k)$ is $x(n)$ delayed over k samples

$x(n+k)$ is $x(n)$ advanced by k samples.

As an example consider a shifted version of the unit impulse. See fig. 1.5. It will be clear that if we multiply an arbitrary signal $x(n)$ with this function we obtain a signal that is zero everywhere except at n = k:

$$y(n) = x(n)\,.\,\delta(n-k) = x(k)\,\delta(n-k)$$

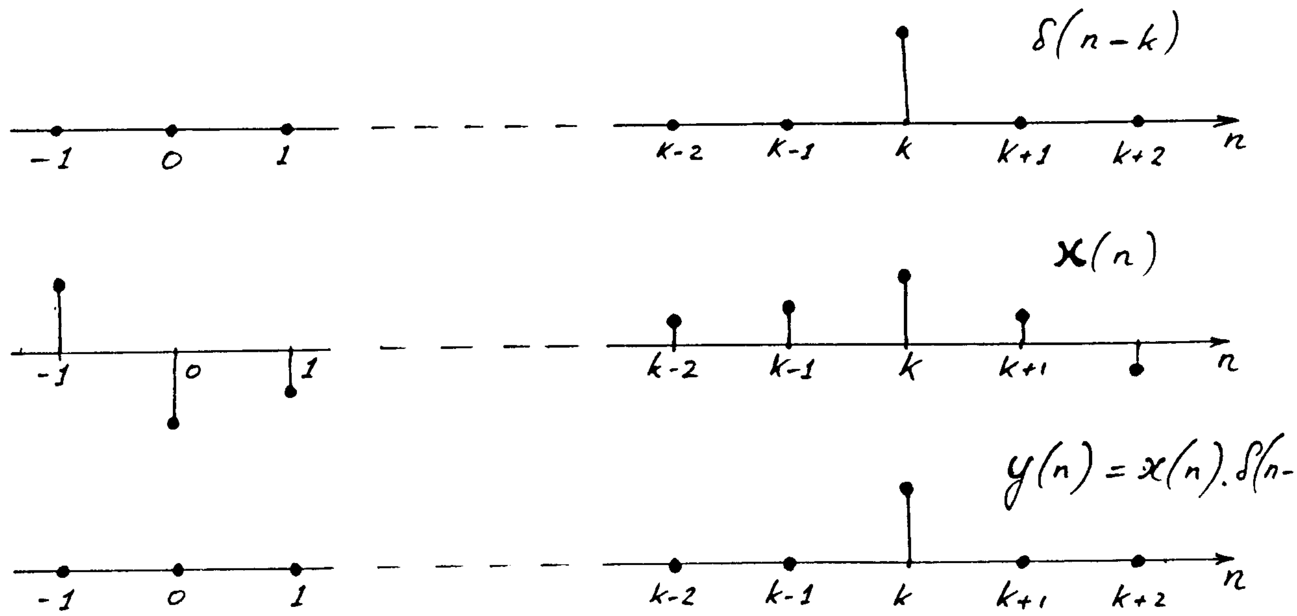

Fig. 1.5.

Doing so for all possible shifts k the original signal can be restored by summing the products so obtained. Thus:

$$x(n) = \sum_{k=-\infty}^{\infty} x(k)\ \delta(n-k) \tag{1.7}$$

Of course if x(n) is of finite duration the infinite sum in (1.7) may be replaced by a finite sum, extending only over the interval where $x(n) \neq 0$.

Eq. (1.7) is a special form of a <u>convolution</u>.
Generally the convolution of two sequences x(n) and y(n) is defined as:

$$x(n) * y(n) = \sum_{k=-\infty}^{\infty} x(k)\ y(n-k) \qquad (1.\text{8}$$

and it is a simple exercise to prove that (commutativity)

$$x(n) * y(n) = y(n) * x(n) = \sum_{k=-\infty}^{\infty} y(k)\ x(n-k) \tag{1.9}$$

Another property of convolution is that it is associative:

$$(x(n) * y(n)) * z(n) = x(n) * (y(n) * z(n))$$
$$= x(n) * y(n) * z(n). \tag{1.10}$$

1.3. Representation in the frequency domain

A frequency domain description, resulting from Fourier transformation is very convenient for analyzing linear systems of all nature, and is therefore quite popular. The conventional Fourier transformation for continuous (time) signals (FTC) will be assumed to be well known.

From the previous paragraphs it can be concluded that a number of fundamental differences exist between discrete-time (or digital) and continuous time (or analogue) signals. Notably are:

1) a discrete time signal is <u>not defined</u> at instants between two samples.
2) Only (relative) frequencies on a finite interval $0 \leq \theta \leq \pi$ need to be considered.

To obtain a frequency domain, or spectral description we will use the Fourier transform for discrete (time) signals (FTD) which takes account of these peculiar properties of discrete signals.

Let x(n) be a discrete time signal. Its Fourier transform, also designated as spectrum, is defined by:

$$X(\theta) = \sum_{n=-\infty}^{\infty} x(n)\ e^{-jn\theta} \qquad (1.\,\cdot\cdot)$$

The inverse transform is given by:

$$x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\theta)\, e^{jn\theta}\, d\theta \qquad (1.12)$$

This transform has already been introduced in the semester I course on Signal Analysis (section 9.2) and therefore we will be very brief on it here.
A number of the most important properties are given in table 1.1 (page 1.8).

From (1.11) it follows that the spectrum of a digital signal is periodic in θ with period 2π. This reflects the fact that sinusoidal signals with relative frequencies that are 2π apart are indistinguishable. (The interval is 2π rather than π since

$$\cos\theta = \frac{e^{j\theta} + e^{-j\theta}}{2}$$

and thus if $0 \leqslant \theta \leqslant \pi$ complex functions $e^{j\varphi}$ with $-\pi < \varphi \leqslant \pi$ have to be considered.)

In eq. (1.12) the integration interval is $(-\pi, \pi)$, but any other interval of length 2π could have been taken. Such an interval will be denoted as a <u>fundamental interval</u> of the spectrum.

<u>Note</u>. In general the fundamental interval will coincide with one period of the spectrum. However, we will encounter signals of which the spectrum will have a period less than 2π i.e. $2\pi/k$. In that case the fundamental interval will contain k periods of the spectrum. The integration in the inverse transformation must always be performed over a fundamental interval, however.

A number of examples of discrete signals and their spectra are given in fig. 1.6 where only the fundamental interval $(-\pi, \pi)$ of the spectrum is shown (this convention will be adopted throughout). It may be instructive to try to derive some of these Fourier transform pairs.

Table 1.1

Properties of the FTD

time domain	frequency domain
real : $x(n) = x^*(n)$	"symmetrical": $X(\theta) = X^*(-\theta)$
even : $x(n) = x(-n)$	real : $X(\theta) = x(0) + 2\sum_{n=1}^{\infty} x(n) \cos n\theta$
odd : $x(n) = -x(-n)$	imaginary: $X(\theta) = 2j\sum_{n=1}^{\infty} x(n) \sin n\theta$

operation	time domain	frequency domain
time shift	$x(n-k)$	$X(\theta)\, e^{-jk\theta}$
convolution	$x(n) * y(n)$	$X(\theta) \cdot Y(\theta)$
product	$x(n) \cdot y(n)$	$X(\theta) * Y(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(\xi)\, Y(\theta-\xi)\, d\xi$.
"frequency shift"	$x(n)\, e^{jn\theta_o}$	$X(\theta - \theta_o)$
"modulation"	$x(n) \cdot \cos n\,\theta_o$	$\frac{1}{2}\left[X(\theta - \theta_o) + X(\theta + \theta_o)\right]$
	$x(n) \cdot \sin n\,\theta_o$	$\frac{1}{2j}\left[X(\theta - \theta_o) - X(\theta + \theta_o)\right]$

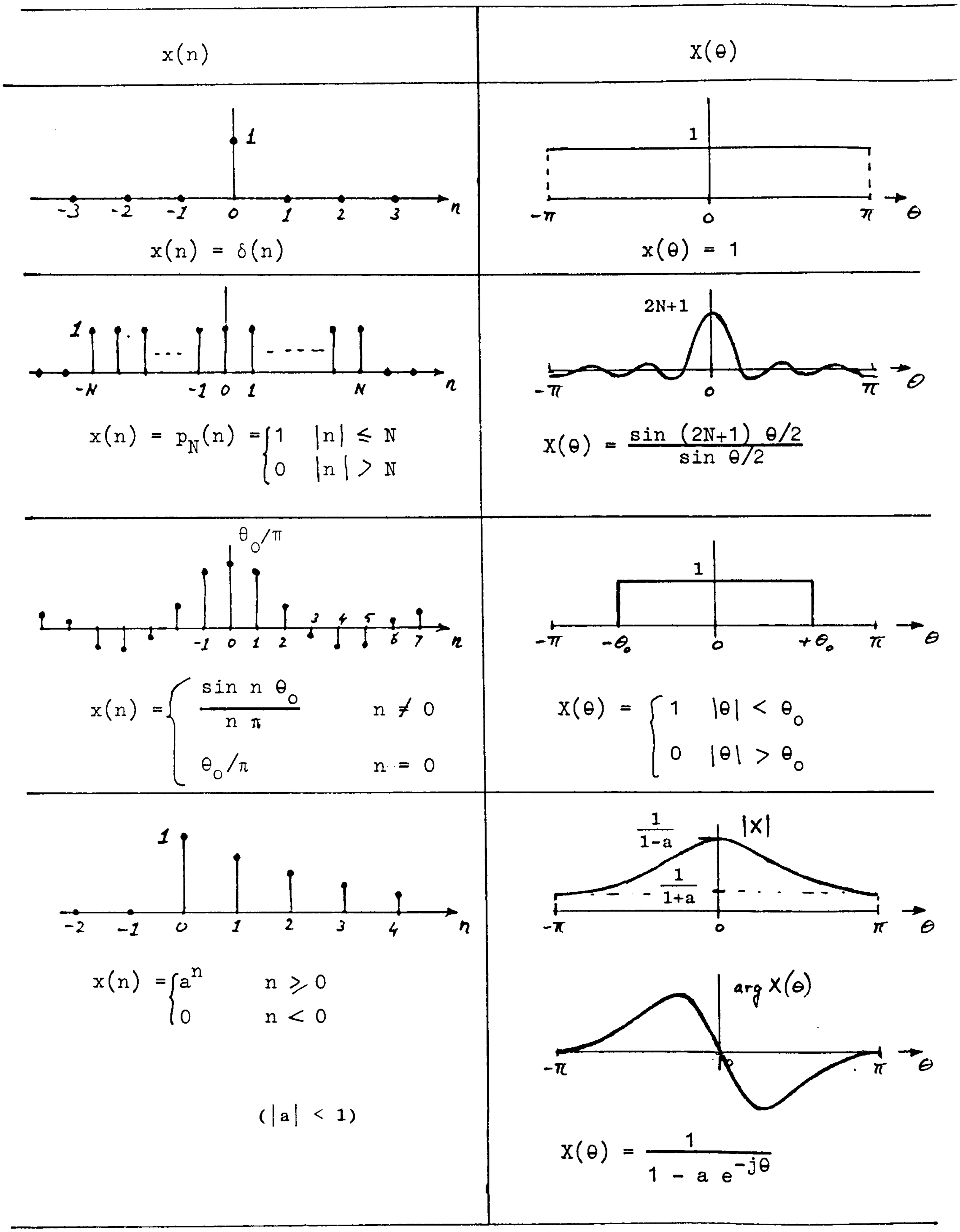
x(n)
X(θ)
x(n) = δ(n)
X(θ) = 1
$x(n) = p_N(n) = \begin{cases} 1 & |n| \leq N \\ 0 & |n| > N \end{cases}$
2N+1
$X(\theta) = \frac{\sin (2N+1)\,\theta/2}{\sin \theta/2}$
θ_0/π
$x(n) = \begin{cases} \frac{\sin n\,\theta_0}{n\,\pi} & n \neq 0 \\ \theta_0/\pi & n = 0 \end{cases}$
$X(\theta) = \begin{cases} 1 & |\theta| < \theta_0 \\ 0 & |\theta| > \theta_0 \end{cases}$
$x(n) = \begin{cases} a^n & n \geq 0 \\ 0 & n < 0 \end{cases}$
(|a| < 1)
$\frac{1}{1-a}$
$\frac{1}{1+a}$
|X|
arg X(θ)
$X(\theta) = \frac{1}{1 - a\,e^{-j\theta}}$

Fig. 1.6.

1.4. The z-transform

For completeness, and because we will need some of its properties later on, we will repeat here the definition of the z-transform. For a more comprehensive treatment see the course notes on Signal Analysis, section 9.

The z-transform of a discrete-time signal x(n) is defined by:

$$\tilde{X}(z) = \sum_{n=-\infty}^{\infty} x(n)\ z^{-n} \tag{1.13}$$

where z is a complex number.

The inverse transform is given by:

$$x(n) = \frac{1}{2\pi j} \oint_C \tilde{X}(z)\ z^{n-1}\ dz \tag{1.14}$$

where C is any closed contour in the region of convergence that encloses z = 0.

Fortunately we will never actually use eq. (1.14), it is merely stated for sake of completeness. In practice we will only encounter z-transforms of rational functions in z, i.e. that are of the form

$$\tilde{X}(z) = \frac{T(z)}{N(z)}$$

where T(z) and N(z) are polynomials in z. For these functions the inverse can be determined without actually needing to perform the contour integration in (1.14). To see this we first determine the z-transform of a number of different signals.

Let

$$x_1(n) = a^n.u(n) \tag{1.15}$$

where u(n) is the unit step function defined in (1.2).
Then from (1.13) it follows:

$$\tilde{X}_1(z) = \sum_{n=-\infty}^{\infty} a^n\ u(n)\ z^{-n} = \sum_{n=0}^{\infty} (a\ z^{-1})^n$$

$$= \frac{1}{1-a\ z^{-1}} \quad \text{for } \left|a\ z^{-1}\right| < 1 \tag{1.16}$$

$$\text{or } |z| > |a|$$

The second example is:

$$x_2(n) = \rho^n \cos n\varphi\ .\ u(n) \tag{1.17}$$

This gives:

$$\tilde{X}_2(z) = \sum_{n=0}^{\infty} \varrho^n \cos n\varphi\, z^{-n} = \frac{1}{2}\sum_{n=0}^{\infty} (\varrho^n e^{jn\varphi} + \varrho^n e^{-jn\varphi})\, z^{-n}$$

$$= \frac{1/2}{1-\varrho e^{j\varphi} z^{-1}} + \frac{1/2}{1-\varrho e^{-j\varphi} z^{-1}}$$

$$= \frac{1 - \varrho \cos\varphi\, z^{-1}}{1 - 2\varrho\cos\varphi\, z^{-1} + \varrho^2 z^{-2}} \quad \text{for } |z| > |\varrho| \tag{1.18}$$

Similarly, the z-transform of

$$x_3(n) = \varrho^n \sin n\varphi \,.\, u(n) \tag{1.19}$$

is:

$$\tilde{X}_3(z) = \frac{\varrho \sin\varphi\, z^{-1}}{1 - 2\varrho\cos\varphi\, z^{-1} + \varrho^2 z^{-2}} \quad \text{for } |z| > |\varrho| \tag{1.20}$$

Now if we have a function of the form

$$\tilde{X}(z) = \frac{a_o + a_1 z^{-1}}{1 - b_1 z^{-1} + b_2 z^{-2}} \tag{1.21}$$

we can rewrite this equation as follows:

$$\tilde{X}(z) = a_o \cdot \frac{1 - \frac{b_1}{2} z^{-1}}{1 - b_1 z^{-1} + b_2 z^{-2}} + \frac{(a_1 + a_o \frac{b_1}{2})\, z^{-1}}{1 - b_1 z^{-1} + b_2 z^{-2}}$$

and from inspection of (1.18) and (1.20) it then follows that

$$x(n) = \left[a_o \varrho^n \cos n\varphi + \left(\frac{a_1 + a_o b_1/2}{\varrho \sin\varphi}\right) \varrho^n \sin n\varphi \right] u(n)$$

where $\varrho = \sqrt{b_2}$.

$$\varphi = \arccos\left(\frac{b_1}{2\sqrt{b_2}}\right)$$

All rational functions that we will encounter can be written in the form

$$\tilde{X}(z) = \sum_{i=1}^{I} \frac{\alpha_i}{1 - \beta_i z^{-1}} + \sum_{j=1}^{J} \frac{a_{oj} + a_{1j} z^{-1}}{1 - b_{1j} z^{-1} + b_{2j} z^{-2}}$$

and thus the inverse can easily be determined using the relations (1.15) - (1.20).

As was already noted in the course on **Signal Analysis**, for those sequences for which the z-transform converges for $|z| = 1$ we have the identity:

$$\tilde{X}(e^{j\theta}) = X(\theta) \tag{1.22}$$

which means that the Fourier transform equals the z-transform when evaluated on the unit circle.