2

Two-Dimensional Systems and Mathematical Preliminaries

2.1 INTRODUCTION

In this chapter we define our notation and discuss some mathematical preliminaries that will be useful throughout the book. Because images are generally outputs of two-dimensional systems, mathematical concepts used in the study of such systems are needed. We start by defining our notation and then review the definitions and properties of linear systems and the Fourier and Z-transforms. This is followed by a review of several fundamental results from matrix theory that are important in digital image processing theory. Two-dimensional random fields and some important concepts from probability and estimation theory are then reviewed. The emphasis is on the final results and their applications in image processing. It is assumed that the reader has encountered most of these basic concepts earlier. The summary discussion provided here is intended to serve as an easy reference for subsequent chapters. The problems at the end of the chapter provide an opportunity to revise these concepts through special cases and examples.

2.2 NOTATION AND DEFINITIONS

A one-dimensional continuous signal will be represented as a function of one variable: $f(x)$, $u(x)$, $s(t)$, and so on. One-dimensional sampled signals will be written as single index sequences: u_n, $u(n)$, and the like.

A continuous image will be represented as a function of two independent variables: $u(x, y)$, $v(x, y)$, $f(x, y)$, and so forth. A sampled image will be represented as a two- (or higher) dimensional sequence of real numbers: $u_{m,n}$, $v(m, n)$, $u(i, j, k)$, and so on. Unless stated otherwise, the symbols $i, j, k, l, m, n, \ldots$ will be used to specify integer indices of arrays and vectors. The symbol roman j will represent $\sqrt{-1}$. The complex conjugate of a complex variable such as z, will be denoted by z^*. Certain symbols will be redefined at appropriate places in the text to keep the notation clear.

Table 2.1 lists several well-known one-dimensional functions that will be often encountered. Their two-dimensional versions are functions of the *separable form*

$$f(x, y) = f_1(x)f_2(y) \tag{2.1}$$

For example, the two-dimensional delta functions are defined as

$$\text{Dirac:} \quad \delta(x, y) = \delta(x)\delta(y) \tag{2.2a}$$

$$\text{Kronecker:} \quad \delta(m, n) = \delta(m)\delta(n) \tag{2.2b}$$

which satisfy the properties

$$\left.\begin{aligned} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x', y')\delta(x - x', y - y')dx'\,dy' = f(x, y) \\ \lim_{\epsilon \to 0}\int_{-\epsilon}^{\epsilon}\int_{-\epsilon}^{\epsilon} \delta(x, y)dx\,dy = 1, \end{aligned}\right\} \tag{2.3}$$

$$\left.\begin{aligned} x(m, n) = \sum\sum_{m', n' = -\infty}^{\infty} x(m', n')\delta(m - m', n - n') \\ \sum\sum_{m, n = -\infty}^{\infty} \delta(m, n) = 1 \end{aligned}\right\} \tag{2.4}$$

The definitions and properties of the functions rect(x, y), sinc(x, y), and comb(x, y) can be defined in a similar manner.

TABLE 2.1 Some Special Functions

Function	Definition	Function	Definition
Dirac delta	$\delta(x) = 0, x \neq 0$	*Rectangle*	$\text{rect}(x) = \begin{cases} 1, & \lvert x \rvert \leq \frac{1}{2} \\ 0, & \lvert x \rvert > \frac{1}{2} \end{cases}$
	$\lim_{\epsilon \to 0} \int_{-\epsilon}^{\epsilon} \delta(x)\, dx = 1$	*Signum*	$\text{sgn}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases}$
Sifting property	$\int_{-\infty}^{\infty} f(x')\delta(x - x')\, dx' = f(x)$	*Sinc*	$\text{sinc}(x) = \dfrac{\sin \pi x}{\pi x}$
Scaling property	$\delta(ax) = \dfrac{\delta(x)}{\lvert a \rvert}$	*Comb*	$\text{comb}(x) = \sum_{n=-\infty}^{\infty} \delta(x - n)$
Kronecker delta	$\delta(n) = \begin{cases} 0, & n \neq 0 \\ 1, & n = 0 \end{cases}$	*Triangle*	$\text{tri}(x) = \begin{cases} 1 - \lvert x \rvert, & \lvert x \rvert \leq 1 \\ 0, & \lvert x \rvert > 1 \end{cases}$
Sifting property	$\sum_{m=-\infty}^{\infty} f(m)\delta(n - m) = f(n)$		

2.3 LINEAR SYSTEMS AND SHIFT INVARIANCE

A large number of imaging systems can be modeled as two-dimensional linear systems. Let $x(m, n)$ and $y(m, n)$ represent the input and output sequences, respectively, of a two-dimensional system (Fig. 2.1), written as

$$y(m, n) = \mathcal{H}[x(m, n)] \tag{2.5}$$

This system is called *linear* if and only if any linear combination of two inputs $x_1(m, n)$ and $x_2(m, n)$ produces the same combination of their respective outputs $y_1(m, n)$ and $y_2(m, n)$, i.e., for arbitrary constants a_1 and a_2

$$\begin{aligned} \mathcal{H}[a_1 x_1(m, n) + a_2 x_2(m, n)] &= a_1 \mathcal{H}[x_1(m, n)] + a_2 \mathcal{H}[x_2(m, n)] \\ &= a_1 y_1(m, n) + a_2 y_2(m, n) \end{aligned} \tag{2.6}$$

This is called *linear superposition.* When the input is the two-dimensional Kronecker delta function at location (m', n'), the output at location (m, n) is defined as

$$h(m, n; m', n') \triangleq \mathcal{H}[\delta(m - m', n - n')] \tag{2.7}$$

and is called the *impulse response* of the system. For an imaging system, it is the image in the output plane due to an ideal point source at location (m', n') in the input plane. In our notation, the semicolon (;) is employed to distinguish the input and output pairs of coordinates.

The impulse response is called the *point spread function* (PSF) when the inputs and outputs represent a positive quantity such as the intensity of light in imaging systems. The term *impulse response* is more general and is allowed to take negative as well as complex values. The *region of support* of an impulse response is the smallest closed region in the m, n plane outside which the impulse response is zero. A system is said to be a *finite impulse response* (FIR) or an *infinite impulse response* (IIR) system if its impulse response has finite or infinite regions of support, respectively.

The output of any linear system can be obtained from its impulse response and the input by applying the superposition rule of (2.6) to the representation of (2.4) as follows:

$$\begin{aligned} y(m, n) &= \mathcal{H}[x(m, n)] \\ &= \mathcal{H}\left[\sum_{m'}\sum_{n'} x(m', n')\delta(m - m', n - n')\right] \\ &= \sum_{m'}\sum_{n'} x(m', n')\mathcal{H}[\delta(m - m', n - n')] \\ \Rightarrow \qquad y(m, n) &= \sum_{m'}\sum_{n'} x(m', n')h(m, n; m', n') \end{aligned} \tag{2.8}$$

$$x(m, n) \longrightarrow \boxed{\mathcal{H}[\cdot]} \longrightarrow y(m, n)$$

Figure 2.1 A system.

A system is called *spatially invariant* or *shift invariant* if a translation of the input causes a translation of the output. Following the definition of (2.7), if the impulse occurs at the origin we will have

$$\mathcal{H}[\delta(m,n)] = h(m,n;0,0)$$

Hence, it must be true for shift invariant systems that

$$\begin{aligned} h(m,n;m',n') &\stackrel{\Delta}{=} \mathcal{H}[\delta(m-m',n-n')] \\ &= h(m-m',n-n';0,0) \\ \Rightarrow \quad h(m,n;m',n') &= h(m-m',n-n') \end{aligned} \tag{2.9}$$

i.e., the impulse response is a function of the two displacement variables only. This means the shape of the impulse response does not change as the impulse moves about the m, n plane. A system is called *spatially varying* when (2.9) does not hold. Figure 2.2 shows examples of PSFs of imaging systems with separable or circularly symmetric impulse responses.

For shift invariant systems, the output becomes

$$y(m,n) = \sum\sum_{m',n'=-\infty}^{\infty} h(m-m',n-n')x(m',n') \tag{2.10}$$

which is called the *convolution* of the input with the *impulse response.* Figure 2.3 shows a graphical interpretation of this operation. The impulse response array is rotated about the origin by 180° and then shifted by (m, n) and overlayed on the array $x(m',n')$. The sum of the product of the arrays $\{x(\cdot,\cdot)\}$ and $\{h(\cdot,\cdot)\}$ in the overlapping regions gives the result at (m, n). We will use the symbol $\circledast$ to denote the convolution operation in both discrete and continuous cases, i.e.,

$$\begin{aligned} g(x,y) &= h(x,y) \circledast f(x,y) \stackrel{\Delta}{=} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x-x',y-y')f(x',y')dx'\,dy' \\ y(m,n) &= h(m,n) \circledast x(m,n) \stackrel{\Delta}{=} \sum\sum_{m'n'=-\infty}^{\infty} h(m-m',n-n')x(m',n') \end{aligned} \tag{2.11}$$

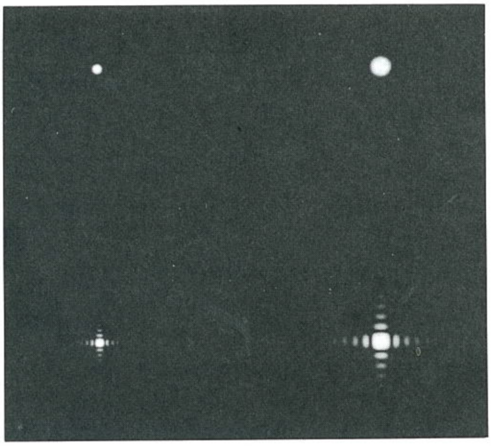
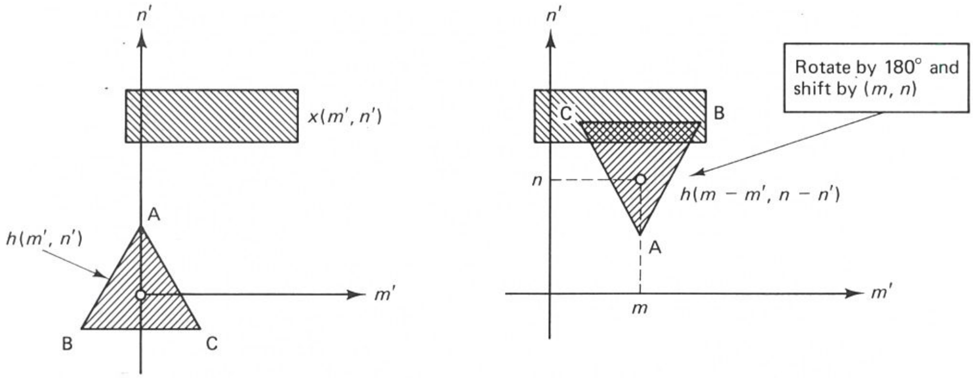


Figure 2.2 Examples of PSFs $\begin{array}{|c|c|}\hline a & b \\ \hline c & d \\ \hline\end{array}$.
(a) Circularly symmetric PSF of average atmospheric turbulence causing small blur; (b) atmospheric turbulence PSF causing large blur; (c) separable PSF of a diffraction limited system with square aperature; (d) same as (c) but with smaller aperture.



(a) impulse response

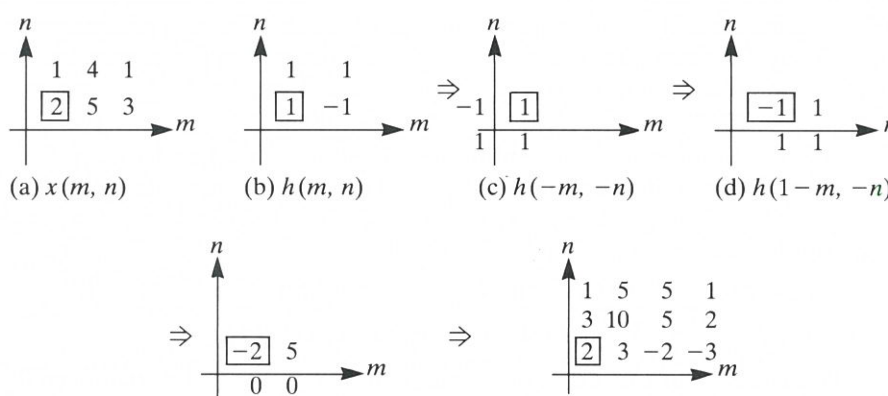
(b) output at location (m, n) is the sum of product of quantities in the area of overlap.

Figure 2.3 Discrete convolution in two dimensions

The convolution operation has several interesting properties, which are explored in Problems 2.2 and 2.3.

Example 2.1 (Discrete convolution)

Consider the 2×2 and 3×2 arrays $h(m, n)$ and $x(m, n)$ shown next, where the boxed element is at the origin. Also shown are the various steps for obtaining the convolution of these two arrays. The result $y(m, n)$ is a 4×3 array. In general, the convolution of two arrays of sizes $(M_1 \times N_1)$ and $(M_2 \times N_2)$ yields an array of size $[(M_1 + M_2 - 1) \times (N_1 + N_2 - 1)]$ (Problem 2.5).



(a) $x(m, n)$ (b) $h(m, n)$ (c) $h(-m, -n)$ (d) $h(1-m, -n)$

(e) $y(1, 0) = -2 + 5 = 3$ (f) $y(m, n)$

2.4 THE FOURIER TRANSFORM

Two-dimensional transforms such as the Fourier transform and the Z-transform are of fundamental importance in digital image processing as will become evident in the subsequent chapters. In one dimension, the Fourier transform of a complex function $f(x)$ is defined as

$$F(\xi) \triangleq \mathcal{F}[f(x)] \triangleq \int_{-\infty}^{\infty} f(x) \exp(-j2\pi\xi x)\, dx \tag{2.12}$$

The inverse Fourier transform of $F(\xi)$ is

$$f(x) \triangleq \mathcal{F}^{-1}[F(\xi)] = \int_{-\infty}^{\infty} F(\xi) \exp(j2\pi\xi x)\, d\xi \tag{2.13}$$

Two-dimensional Fourier transform and its inverse are defined analogously by the linear transformations

$$F(\xi_1, \xi_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y) \exp[-j2\pi(x\xi_1 + y\xi_2)]\, dx\, dy \tag{2.14}$$

$$f(x, y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(\xi_1, \xi_2) \exp[j2\pi(x\xi_1 + y\xi_2)]\, d\xi_1\, d\xi_2 \tag{2.15}$$

Examples of some useful two-dimensional Fourier transforms are given in Table 2.2.

Properties of the Fourier Transform

Table 2.3 gives a summary of the properties of the two-dimensional Fourier transform. Some of these properties are discussed next.

1. *Spatial frequencies.* If $f(x, y)$ is luminance and x, y the spatial coordinates, then ξ_1, ξ_2 are the spatial frequencies that represent luminance changes with respect to spatial distances. The units of ξ_1 and ξ_2 are reciprocals of x and y, respectively. Sometimes the coordinates x, y are normalized by the viewing distance of the image $f(x, y)$. Then the units of ξ_1, ξ_2 are cycles per degree (of the viewing angle).
2. *Uniqueness.* For continuous functions, $f(x, y)$ and $F(\xi_1, \xi_2)$ are unique with respect to one another. There is no loss of information if instead of preserving the image, its Fourier transform is preserved. This fact has been utilized in an image data compression technique called *transform coding.*
3. *Separability.* By definition, the Fourier transform kernel is separable, so that it

TABLE 2.2 Two-Dimensional Fourier Transform Pairs

$f(x, y)$	$F(\xi_1, \xi_2)$
$\delta(x, y)$	1
$\delta(x \pm x_0, y \pm y_0)$	$\exp(\pm j2\pi x_0\xi_1)\exp(\pm j2\pi y_0\xi_2)$
$\exp(\pm j2\pi x\eta_1)\exp(\pm j2\pi y\eta_2)$	$\delta(\xi_1 \mp \eta_1, \xi_2 \mp \eta_2)$
$\exp[-\pi(x^2 + y^2)]$	$\exp[-\pi(\xi_1^2 + \xi_2^2)]$
$\mathrm{rect}(x, y)$	$\mathrm{sinc}(\xi_1, \xi_2)$
$\mathrm{tri}(x, y)$	$\mathrm{sinc}^2(\xi_1, \xi_2)$
$\mathrm{comb}(x, y)$	$\mathrm{comb}(\xi_1, \xi_2)$

TABLE 2.3 Properties of Two-Dimensional Fourier Transform

Property	Function $f(x, y)$	Fourier Transform $F(\xi_1, \xi_2)$
Rotation	$f(\pm x, \pm y)$	$F(\pm\xi_1, \pm\xi_2)$
Linearity	$a_1 f_1(x, y) + a_2 f_2(x, y)$	$a_1 F_1(\xi_1, \xi_2) + a_2 F_2(\xi_1, \xi_2)$
Conjugation	$f^*(x, y)$	$F^*(-\xi_1, -\xi_2)$
Separability	$f_1(x) f_2(y)$	$F_1(\xi_1) F_2(\xi_2)$
Scaling	$f(ax, by)$	$\dfrac{F(\xi_1/a, \xi_2/b)}{\lvert ab \rvert}$
Shifting	$f(x \pm x_0, y \pm y_0)$	$\exp[\pm j2\pi(x_0 \xi_1 + y_0 \xi_2)] F(\xi_1, \xi_2)$
Modulation	$\exp[\pm j2\pi(\eta_1 x + \eta_2 y)] f(x, y)$	$F(\xi_1 \mp \eta_1, \xi_2 \mp \eta_2)$
Convolution	$g(x, y) = h(x, y) \circledast f(x, y)$	$G(\xi_1, \xi_2) = H(\xi_1, \xi_2) F(\xi_1, \xi_2)$
Multiplication	$g(x, y) = h(x, y) f(x, y)$	$G(\xi_1, \xi_2) = H(\xi_1, \xi_2) \circledast F(\xi_1, \xi_2)$
Spatial correlation	$c(x, y) = h(x, y) \star f(x, y)$	$C(\xi_1, \xi_2) = H(-\xi_1, -\xi_2) F(\xi_1, \xi_2)$
Inner product	$I = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y) h^*(x, y)\, dx\, dy$	$I = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(\xi_1, \xi_2) H^*(\xi_1, \xi_2)\, d\xi_1\, d\xi_2$

can be written as a separable transformation in x and y, i.e.,

$$F(\xi_1, \xi_2) = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(x, y) \exp(-j2\pi x \xi_1)\, dx\right] \exp(-j2\pi y \xi_2)\, dy$$

This means the two-dimensional transformation can be realized by a succession of one-dimensional transformations along each of the spatial coordinates.

4. *Frequency response and eigenfunctions of shift invariant systems.* An eigenfunction of a system is defined as an input function that is reproduced at the output with a possible change only in its amplitude. A fundamental property of a linear shift invariant system is that its eigenfunctions are given by the complex exponential $\exp[j2\pi(\xi_1 x + \xi_2 y)]$. Thus in Fig. 2.4, for any fixed (ξ_1, ξ_2), the output of the linear shift invariant system would be

$$g(x, y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x - x', y - y') \exp[j2\pi(\xi_1 x' + \xi_2 y')]\, dx'\, dy'$$

Performing the change of variables $\tilde{x} = x - x'$, $\tilde{y} = y - y'$ and simplifying the result, we get

$$g(x, y) = H(\xi_1, \xi_2) \exp[j2\pi(\xi_1 x + \xi_2 y)] \tag{2.16}$$

The function $H(\xi_1, \xi_2)$, which is the Fourier transform of the impulse response, is also called the *frequency response* of the system. It represents the (complex) amplitude of the system response at spatial frequency (ξ_1, ξ_2).

$$\phi \longrightarrow \boxed{h(x, y)} \longrightarrow H\phi$$

Figure 2.4 Eigenfunctions of a linear shift invariant system. $\phi \overset{\Delta}{=} \exp\{j2\pi(\xi_1 x + \xi_2 y)\}$, $H = H(\xi_1, \xi_2) \overset{\Delta}{=}$ Fourier transform of $h(x, y)$.

5. *Convolution theorem.* The Fourier transform of the convolution of two functions is the product of their Fourier transforms, i.e.,

$$g(x, y) = h(x, y) \circledast f(x, y) \Leftrightarrow G(\xi_1, \xi_2) = H(\xi_1, \xi_2) F(\xi_1, \xi_2) \tag{2.17}$$

This theorem suggests that the convolution of two functions may be evaluated by inverse Fourier transforming the product of their Fourier transforms. The discrete version of this theorem yields a fast Fourier transform based convolution algorithm (see Chapter 5).

The converse of the convolution theorem is that the Fourier transform of the product of two functions is the convolution of their Fourier transforms.

The result of convolution theorem can also be extended to the *spatial correlation* between two real functions $h(x, y)$ and $f(x, y)$, which is defined as

$$c(x, y) = h(x, y) \bigstar f(x, y) \triangleq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x', y') f(x + x', y + y')\, dx'\, dy' \tag{2.18}$$

A change of variables shows that $c(x, y)$ is also the convolution $h(-x, -y) \circledast f(x, y)$, which yields

$$C(\xi_1, \xi_2) = H(-\xi_1, -\xi_2) F(\xi_1, \xi_2) \tag{2.19}$$

6. *Inner product preservation.* Another important property of the Fourier transform is that the inner product of two functions is equal to the inner product of their Fourier transforms, i.e.,

$$I \triangleq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) h^*(x, y)\, dx\, dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\xi_1, \xi_2) H^*(\xi_1, \xi_2)\, d\xi_1\, d\xi_2 \tag{2.20}$$

Setting $h = f$, we obtain the well-known *Parseval energy conservation formula*

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(x, y)|^2\, dx\, dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |F(\xi_1, \xi_2)|^2\, d\xi_1\, d\xi_2 \tag{2.21}$$

i.e., the total energy in the function is the same as in its Fourier transform.

7. *Hankel transform.* The Fourier transform of a circularly symmetric function is also circularly symmetric and is given by what is called the *Hankel transform* (see Problem 2.10).

Fourier Transform of Sequences (Fourier Series)

For a *one-dimensional* sequence $x(n)$, real or complex, its Fourier transform is defined as the series

$$X(\omega) = \sum_{n=-\infty}^{\infty} x(n) \exp(-jn\omega), \qquad -\pi \le \omega < \pi \tag{2.22}$$

The inverse transform is given by

$$x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\omega) \exp(jn\omega)\, d\omega \tag{2.23}$$

Note that $X(\omega)$ is periodic with period 2π. Hence it is sufficient to specify it over one period.

The Fourier transform pair of a two-dimensional sequence $x(m, n)$ is defined as

$$X(\omega_1, \omega_2) \triangleq \sum_{m,n=-\infty}^{\infty}\sum x(m, n) \exp[-j(m\omega_1 + n\omega_2)], \qquad -\pi \le \omega_1, \omega_2 < \pi \tag{2.24}$$

$$x(m, n) = \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} X(\omega_1, \omega_2) \exp[j(m\omega_1 + n\omega_2)]\, d\omega_1\, d\omega_2 \tag{2.25}$$

Now $X(\omega_1, \omega_2)$ is periodic with period 2π in each argument, i.e.,

$$X(\omega_1 \pm 2\pi, \omega_2 \pm 2\pi) = X(\omega_1 \pm 2\pi, \omega_2) = X(\omega_1, \omega_2 \pm 2\pi) = X(\omega_1, \omega_2) \tag{2.25}$$

Often, the sequence $x(m, n)$ in the series in (2.24) is absolutely summable, i.e.,

$$\sum_{m,n=-\infty}^{\infty}\sum |x(m, n)| < \infty \tag{2.26}$$

Analogous to the continuous case, $H(\omega_1, \omega_2)$, the Fourier transform of the shift invariant impulse response is called *frequency response*. The Fourier transform of sequences has many properties similar to the Fourier transform of continuous functions. These are summarized in Table 2.4.

TABLE 2.4 Properties and Examples of Fourier Transform of Two-Dimensional Sequences

Property	Sequence	Transform
	$x(m, n), y(m, n), h(m, n), \cdots$	$X(\omega_1, \omega_2), Y(\omega_1, \omega_2), H(\omega_1, \omega_2), \cdots$
Linearity	$a_1 x_1(m, n) + a_2 x_2(m, n)$	$a_1 X_1(\omega_1, \omega_2) + a_2 X_2(\omega_1, \omega_2)$
Conjugation	$x^*(m, n)$	$X^*(-\omega_1, -\omega_2)$
Separability	$x_1(m) x_2(n)$	$X_1(\omega_1) X_2(\omega_2)$
Shifting	$x(m \pm m_0, n \pm n_0)$	$\exp[\pm j(m_0\omega_1 + n_0\omega_2)] X(\omega_1, \omega_2)$
Modulation	$\exp[\pm j(\omega_{01} m + \omega_{02} n)] x(m, n)$	$X(\omega_1 \mp \omega_{01}, \omega_2 \mp \omega_{02})$
Convolution	$y(m, n) = h(m, n) \circledast x(m, n)$	$Y(\omega_1, \omega_2) = H(\omega_1, \omega_2) X(\omega_1, \omega_2)$
Multiplication	$h(m, n) x(m, n)$	$\left(\frac{1}{4\pi^2}\right) H(\omega_1, \omega_2) \circledast X(\omega_1, \omega_2)$
Spatial correlation	$c(m, n) = h(m, n) \bigstar x(m, n)$	$C(\omega_1, \omega_2) = H(-\omega_1, -\omega_2) X(\omega_1, \omega_2)$
Inner product	$I = \sum_{m,n=-\infty}^{\infty}\sum x(m, n) y^*(m, n)$	$I = \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} X(\omega_1, \omega_2) Y^*(\omega_1, \omega_2)\, d\omega_1\, d\omega_2$
Energy conservation	$\mathscr{E} = \sum_{m,n=-\infty}^{\infty}\sum \|x(m, n)\|^2$	$\mathscr{E} = \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \|X(\omega_1, \omega_2)\|^2\, d\omega_1\, d\omega_2$
	$\sum_{m,n=-\infty}^{\infty}\sum \exp[j(m\omega_{01} + n\omega_{02})]$	$4\pi^2 \delta(\omega_1 - \omega_{01}, \omega_2 - \omega_{02})$
	$\delta(m, n)$	$\frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \exp[-j(\omega_1 m + \omega_2 n)]\, d\omega_1\, d\omega_2$

2.5 THE Z-TRANSFORM OR LAURENT SERIES

A useful generalization of the Fourier series is the Z-transform, which for a two-dimensional complex sequence $x(m, n)$ is defined as

$$X(z_1, z_2) = \sum\sum_{m,n=-\infty}^{\infty} x(m, n) z_1^{-m} z_2^{-n} \tag{2.27}$$

where z_1, z_2 are complex variables. The set of values of z_1, z_2 for which this series converges uniformly is called the *region of convergence*. The Z-transform of the impulse response of a linear shift invariant discrete system is called its *transfer function*. Applying the convolution theorem for Z-transforms (Table 2.5) we can transform (2.10) as

$$Y(z_1, z_2) = H(z_1, z_2) X(z_1, z_2)$$

$$\Rightarrow \quad H(z_1, z_2) = \frac{Y(z_1, z_2)}{X(z_1, z_2)}$$

i.e., the transfer function is also the ratio of the Z-transforms of the output and the input sequences. The inverse Z-transform is given by the double contour integral

$$x(m, n) = \frac{1}{(j2\pi)^2} \oint\oint X(z_1, z_2) z_1^{m-1} z_2^{n-1} \, dz_1 \, dz_2 \tag{2.28}$$

where the contours of integration are counterclockwise and lie in the region of convergence. When the region of convergence includes the unit circles $|z_1| = 1, |z_2| = 1$, then evaluation of $X(z_1, z_2)$ at $z_1 = \exp(j\omega_1), z_2 = \exp(j\omega_2)$ yields the Fourier transform of $x(m, n)$. Sometimes $X(z_1, z_2)$ is available as a finite series (such as the transfer function of FIR filters). Then $x(m, n)$ can be obtained by inspection as the coefficient of the term $z_1^{-m} z_2^{-n}$.

TABLE 2.5 Properties of the Two-Dimensional Z-Transform

Property	Sequence	Z-Transform
	$x(m, n), y(m, n), h(m, n), \cdots$	$X(z_1, z_2), Y(z_1, z_2), H(z_1, z_2), \cdots$
Rotation	$x(-m, -n)$	$X(z_1^{-1}, z_2^{-1})$
Linearity	$a_1 x_1(m, n) + a_2 x_2(m, n)$	$a_1 X_1(z_1, z_2) + a_2 X_2(z_1, z_2)$
Conjugation	$X^*(m, n)$	$X^*(z_1^*, z_2^*)$
Separability	$x_1(m) x_2(n)$	$X_1(z_1) X_2(z_2)$
Shifting	$x(m \pm m_0, n \pm n_0)$	$z_1^{\pm m_0} z_2^{\pm n_0} X(z_1, z_2)$
Modulation	$a^m b^n x(m, n)$	$X\left(\frac{z_1}{a}, \frac{z_2}{b}\right)$
Convolution	$h(m, n) \circledast x(m, n)$	$H(z_1, z_2) X(z_1, z_2)$
Multiplication	$x(m, n) y(m, n)$	$\left(\frac{1}{2\pi j}\right)^2 \oint\oint_{C_1 C_2} X\left(\frac{z_1}{z_1'}, \frac{z_2}{z_2'}\right) Y(z_1', z_2') \frac{dz_1'}{z_1'} \frac{dz_2'}{z_2'}$

Causality and Stability

A one-dimensional shift invariant system is called causal if its output at any time is not affected by future inputs. This means its impulse response $h(n) = 0$ for $n < 0$ and its transfer function must have a one-sided Laurent series, i.e.,

$$H(z) = \sum_{n=0}^{\infty} h(n) z^{-n} \tag{2.29}$$

Extending this definition, any sequence $x(n)$ is called *causal* if $x(n) = 0, n < 0$; *anticausal* if $x(n) = 0, n \geq 0$, and *noncausal* if it is neither causal nor anticausal.

A system is called *stable* if its output remains uniformly bounded for any bounded input. For linear shift invariant systems, this condition requires that the impulse response should be absolutely summable (prove it!), i.e.,

$$\sum_{n=-\infty}^{\infty} |h(n)| < \infty \tag{2.30}$$

This means $H(z)$ cannot have any poles on the unit circle $|z| = 1$. If this system is to be *causal and stable,* then the convergence of (2.29) at $|z| = 1$ implies the series must converge for all $|z| \geq 1$, i.e., the poles of $H(z)$ must lie *inside* the unit circle.

In two dimensions, a linear shift invariant system is stable when

$$\sum_{m} \sum_{n} |h(m, n)| < \infty \tag{2.31}$$

which implies the region of convergence of $H(z_1, z_2)$ must include the unit circles, i.e., $|z_1| = 1, |z_2| = 1$.

2.6 OPTICAL AND MODULATION TRANSFER FUNCTIONS

For a spatially invariant imaging system, its *optical transfer function* (OTF) is defined as its normalized frequency response, i.e.,

$$\text{OTF} = \frac{H(\xi_1, \xi_2)}{H(0, 0)} \tag{2.32}$$

The *modulation transfer function* (MTF) is defined as the magnitude of the OTF, i.e.,

$$\text{MTF} = |\text{OTF}| = \frac{|H(\xi_1, \xi_2)|}{|H(0, 0)|} \tag{2.33}$$

Similar relations are valid for discrete systems. Figure 2.5 shows the MTFs of systems whose PSFs are displayed in Fig. 2.2. In practice, it is often the MTF that is measurable. The phase of the frequency response is estimated from physical considerations. For many optical systems, the OTF itself is positive.