Chapter 4
The Gabor Transform and Time–Frequency Signal Analysis

What is clear and easy to grasp attracts us; complications deter.

David Hilbert

Motivated by 'quantum mechanics', in 1946 the physicist Gabor defined elementary time-frequency atoms as waveforms that have a minimal spread in a time-frequency plane. To measure time-frequency 'information' content, he proposed decomposing signals over these elementary atomic waveforms. By showing that such decompositions are closely related to our sensitivity to sounds, and that they exhibit important structures in speech and music recordings, Gabor demonstrated the importance of localized time-frequency signal processing.

Stéphane Mallat

4.1 Introduction

Signals are, in general, nonstationary. A complete representation of nonstationary signals requires frequency analysis that is local in time, resulting in the time–frequency analysis of signals. The Fourier transform analysis has long been recognized as the great tool for the study of stationary signals and processes where the properties are statistically invariant over time. However, it cannot be used for the frequency analysis that is local in time. In recent years, several useful methods have been developed for the time–frequency signal analysis. They include the Gabor transform, the Zak transform, and the wavelet transform.

It has already been stated in Sect. 1.2 that decomposition of a signal into a small number of elementary waveforms that are localized in time and frequency plays a remarkable role in signal processing. Such a decomposition reveals important structures in analyzing nonstationary signals such as speech and music. In order to measure localized frequency components of sounds, Gabor (1946) first introduced the windowed Fourier transform (or the local time–frequency transform), which may be called the Gabor transform, and suggested the representation of a signal in a joint time–frequency domain. Subsequently, the Gabor transform analysis has effectively been applied in many fields of science and engineering, such as image

L. Debnath, F.A. Shah, *Wavelet Transforms and Their Applications*,
DOI 10.1007/978-0-8176-8418-1_4

analysis and image compression, object and pattern recognition, computer vision, optics, and filter banks. Since medical signal analysis and medical signal processing play a crucial role in medical diagnostics, the Gabor transform has also been used for the study of brain functions, ECC signals, and other medical signals.

This chapter deals with classification of signals, joint time–frequency analysis of signals, and the Gabor transform and its basic properties, including the inversion formula. Special attention is given to the discrete Gabor transform and the Gabor representation problem. Included are the Zak transform, its basic properties, and applications for studying the orthogonality and completeness of the Gabor frames in the critical case.

4.2 Classification of Signals and the Joint Time–Frequency Analysis of Signals

Many physical quantities including pressure, sound waves, electric fields, voltage, electric current, and electromagnetic fields vary with time t. These quantities are called *signals* or *waveforms*. Example of signals include speech signals, optical signals, acoustic signals, biomedical signals, radar, and sonar. Indeed, signals are very common in the real world.

In general, there are two kinds of signals: (a) *deterministic* and (b) *random* (or *stochastic*). A signal is called *deterministic* if it can be determined explicitly, under identical conditions, in terms of a mathematical relationship. A deterministic signal is referred to as *periodic* or *transient* if the signal repeats continuously at regular intervals of time or decays to zero after a finite time interval. Periodic and transient signals are shown in Figs. 4.1a, b and 4.2.

On the other hand, signals are, in general, random or stochastic in nature in the sense that they cannot be determined precisely at any given instant of time even under identical conditions. Obviously, probabilistic and statistical information is required for a description of random signals. It is necessary to consider a particular random process that can produce a set of time-histories, known as an *ensemble*. This can represent an experiment producing random data, which is repeated n times to give an ensemble of n separate records (see Fig. 4.3).

The *average value* at time t over the ensemble x is defined by

$$\langle x(t) \rangle = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} x_k(t), \tag{4.2.1}$$

where x takes any one of a set of values x_k, and $k = 1, 2, \ldots, n$.

The average value of the product of two samples taken at two separate times t_1 and t_2 is called the *autocorrelation function* R, for each separate record, defined by

$$R(\tau) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} x_k(t_1)\, x_k(t_2), \tag{4.2.2}$$

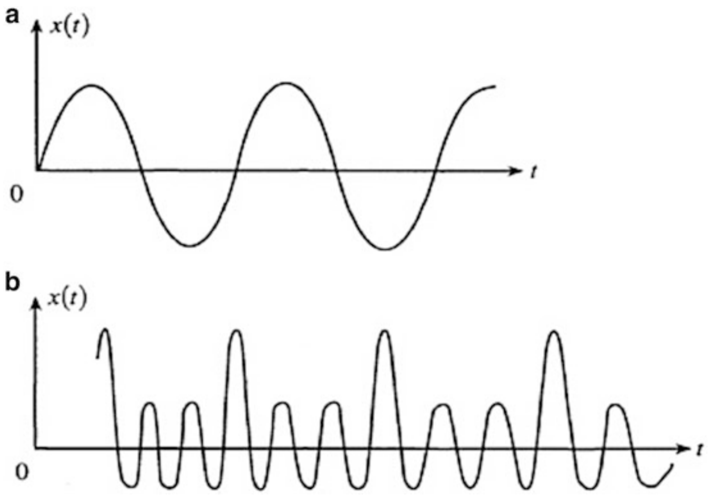

Fig. 4.1 (**a**) Sinusoidal periodic signal; (**b**) nonsinusoidal periodic signal

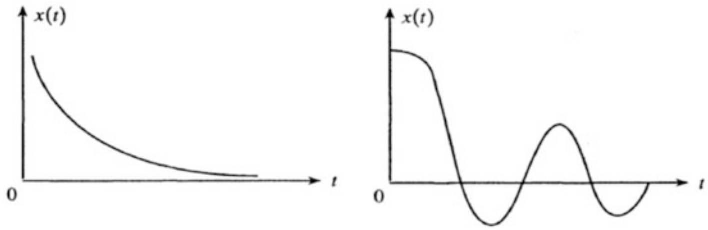

Fig. 4.2 Transient signals

where $\tau = t_1 - t_2$. The process of finding these values is referred to as *ensemble averaging* and may be continued over the entire record length to provide statistical information on the complex set of records.

A signal is called *stationary* if the values of $\langle x(t) \rangle$ and $R(t)$ remain constant for all possible values of t and $R(\tau)$ depends only on the time displacement $\tau = t_1 - t_2$ (see Fig. 4.4a). In most practical situations, a signal is called stationary if $\langle x(t) \rangle$ and $R(\tau)$ are constant over the finite record length T.

A signal is called *nonstationary* if the values of $\langle x(t) \rangle$ and $R(\tau)$ vary with time (see Fig. 4.4b). However, in many practical situations, the change of time is very slow, so the signal can be regarded as stationary. Under certain conditions, we regard a signal as stationary by considering the statistical characteristic of a single long record. The average value of a signal $x(t)$ over a time length T is defined by

$$\bar{x} = \lim_{T \to \infty} \frac{1}{T} \int_0^T x(t)\, dt, \tag{4.2.3}$$

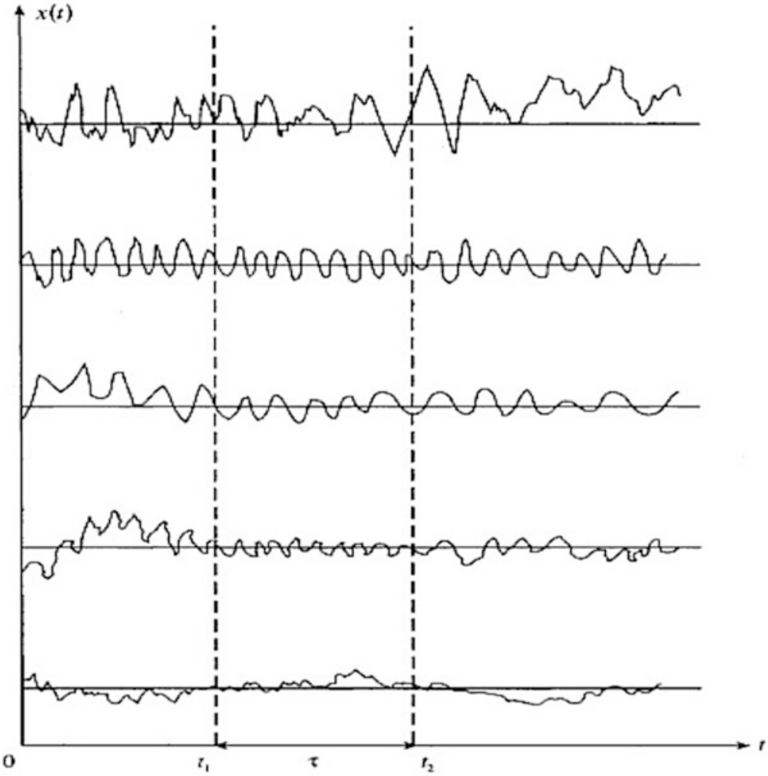

Fig. 4.3 Ensemble of n records

where $\bar{x}$ is used to represent a single time-history average to distinguish it from the ensemble average $\langle x \rangle$.

Similarly, the *autocorrelation function* over a single time length T is defined by

$$R(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T x(t)\, x(t+\tau)\, dt. \tag{4.2.4}$$

Under certain circumstances, the ensemble average can be obtained from computing the time average so

$$\langle x \rangle = \bar{x} \tag{4.2.5}$$

for all values of time t. Then, this process is called an *ergodic random process*. By definition, this must be a stationary process. However, the converse is not necessarily true, that is, a stationary random process need not be ergodic.

Finally, we can introduce various ensemble averages of $x(t)$ which take any one of the values $x_k(t), k = 1, 2, \ldots, n$ at time t in terms of probability $P_x\big(x_k(t)\big)$. The ensemble average of x is then defined by

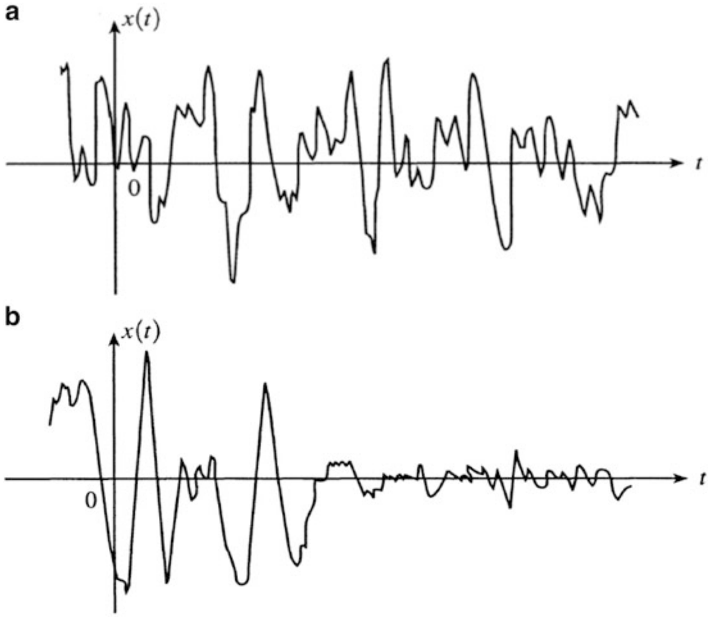

Fig. 4.4 (**a**) Stationary random signal. (**b**) Nonstationary random signal

$$\langle x \rangle = \sum_{k=1}^{n} P_x\left(x_k\right) x_k. \tag{4.2.6}$$

We now consider two random variables $x_i(t)$ and $x_k(s)$ which are values of a random process x at times t and s with the joint probability distribution $P_x^{t,s}\left(x_i, x_k\right)$. Then, the *autocorrelation function*, $R(t,s)$ of the random process x is defined by

$$R(t,s) = \left\langle x(t)\, x(s) \right\rangle = \sum_{i,k} P_x^{t,s}\left(x_i, x_k\right) x_i\, x_k. \tag{4.2.7}$$

This function provides a great deal of information about the random process and arises often in signal analysis. For a random stationary process, $P_x^{(t,s)}$ is a function of $\tau = t - s$ only, so that

$$R(t,s) = R(t-s) = R(\tau) \tag{4.2.8}$$

and hence, $R(-\tau) = R(\tau)$ and R is an even function.

Signals can be described in a time domain or in a frequency domain by the traditional method of Fourier transform analysis. The frequency description of signals is known as the *frequency* (or *spectral*) analysis. It was recognized long ago that a global Fourier transform of a long time signal is of little practical value in analyzing the frequency spectrum of the signal. From the Fourier spectrum (or spectral function) $\hat{f}(\omega)$ of a signal $f(t)$, it is always possible to determine which frequencies were present in the signal. However, there is absolutely no indication as

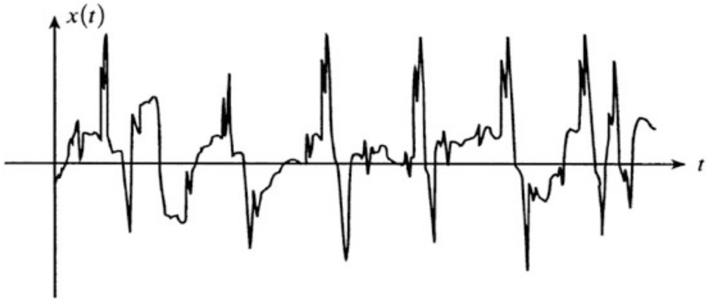

Fig. 4.5 ECG signal of a human heart

to when those frequencies existed. So, the Fourier transform analysis cannot provide any information regarding either a time evolution of spectral characteristics or a possible localization with respect to the time variable. Transient signals such as a speech signals or ECG signals (see Fig. 4.5) require the idea of frequency analysis that is local in time.

In general, the frequency of a signal varies with time, so there is a need for a joint time–frequency representation of a signal in order to describe fully the characteristics of the signal. Thus, both the analysis and processing of nonstationary signals require specific mathematical methods which go beyond the classical Fourier transform analysis. Gabor (1946) was the first to introduce the joint time–frequency representation of a signal. Almost simultaneously, Ville (1948) first introduced the Wigner distribution into time–frequency signal analysis to unfold the signal in the time–frequency plane in such a way that this development led to a joint representation in time–frequency atoms.

4.3 Definition and Examples of the Gabor Transform

Gabor (1946) first introduced a time-localization window function $g_a(t-b)$ for extracting local information from a Fourier transform of a signal, where the parameter a measures the width of the window, and the parameter b is used to represent translation of the window to cover the whole time domain. The idea is to use this window function in order to localize the Fourier transform, then shift the window to another position, and so on. This remarkable property of the Gabor transform provides the local aspect of the Fourier transform with time resolution equal to the size of the window. Thus, the Gabor transform is often called the *windowed Fourier transform*. Gabor first introduced

$$g_{t,\omega}(\tau) = \exp(i\,\omega\tau)\, g(\tau - t) = M_\omega\, T_t\, g(\tau), \tag{4.3.1}$$

as the window function by first translating in time and then modulating the function $g(t) = \pi^{-\frac{1}{4}} \exp\left(-2^{-1}t^2\right)$ which is called the *canonical coherent state* in quantum physics. The energy associated with the function $g_{t,\omega}$ is localized in the neighborhood of t in an interval of size σ_t, measured by the standard deviation of $|g|^2$. Evidently, the Fourier transform of $g_{t,\omega}(\tau)$ with respect to τ is given by

$$\hat{g}_{t,\omega}(\upsilon) = \hat{g}(\upsilon - \omega) \exp\left\{ -it(\upsilon - \omega)\right\}. \tag{4.3.2}$$

Obviously, the energy of $\hat{g}_{t,\omega}$ is concentrated near the frequency ω in an interval of size σ_ω which measures the frequency dispersion (or bandwidth) of $\hat{g}_{t,\omega}$. In a time–frequency (t, ω) plane, the energy spread of the Gabor atom $\hat{g}_{t,\omega}$ can be represented by the rectangle with center at $\left(\langle t\rangle, \langle\omega\rangle\right)$ and sides σ_t (along the time axis) and σ_ω (along the frequency axis). According to the Heisenberg uncertainty principle, the area of the rectangle is at least $\frac{1}{2}$; that is, $\sigma_t\, \sigma_\omega \geq \frac{1}{2}$. This area is minimum when g is a Gaussian function, and the corresponding $g_{t,\omega}$ is called the *Gabor function* (or *Gabor wavelet*).

Definition 4.3.1 (The Continuous Gabor Transform). The continuous Gabor transform of a function $f \in L^2(\mathbb{R})$ with respect to a window function $g \in L^2(\mathbb{R})$ is denoted by $\mathscr{G}[f](t, \omega) = \tilde{f}_g(t, \omega)$ and defined by

$$\mathscr{G}[f](t, \omega) = \tilde{f}_g(t, \omega) = \int_{-\infty}^{\infty} f(\tau)\, g(\tau - t)\, e^{-i\omega\tau} d\tau = \left\langle f, \overline{g_{t,\omega}}\right\rangle, \tag{4.3.3}$$

where $g_{t,\omega}(\tau) = \overline{g}(\tau - t) \exp(i\omega\tau)$, so, $\left\|g_{t,\omega}\right\| = \left\|g\right\|$ and hence, $g_{t,\omega} \in L^2(\mathbb{R})$.

Clearly, the Gabor transform $\tilde{f}_g(t, \omega)$ of a given signal f depends on both time t and frequency ω. For any fixed t, $\tilde{f}_g(t, \omega)$ represents the frequency distribution at time t. Usually, only values of $f(\tau)$ for $\tau \leq t$ can be used in computing $\tilde{f}_g(t, \omega)$. In a system of finite memory, there exists a time interval $T > 0$ such that only the values $f(\tau)$ for $\tau > t - T$ can affect the output at time t. Thus, the transform function $\tilde{f}_g(t, \omega)$ depends only on $f(\tau)$ for $t - T \leq \tau \leq t$. Mathematically, if $g_{t,\omega}(\tau)$ vanishes outside $[-T, 0]$ such that $\operatorname{supp} g \subset [-T, 0]$, then $g_{t,\omega}(\tau)$ can be used to localize the signal in time. For any $t \in \mathbb{R}$, we can define $f_t(\tau) = g(\tau - t) f(\tau)$ so that $\operatorname{supp} f_t \subset [t - T, t]$. Therefore, $f_t(\tau)$ can be regarded as a localized version of f that depends only on the values of $f(\tau)$ in $t - T \leq \tau \leq t$. If g is continuous, then the values of $f_t(\tau)$ with $\tau \approx t - T$ and $\tau \approx t$ are small. This means that the localization is smooth, and this particular feature plays an important role in signal processing.

In physical applications, f and g represent signals with finite energy. In quantum physics, $\tilde{f}_g(t, \omega)$ is referred to as the *canonical coherent state* representation of f. The term *coherent state* was first used by Glauber (1964) in quantum optics.

We next discuss the following consequences of the preceding definition.

1. For a fixed t, the Fourier transform of $f_t(\tau)$ with respect to τ is given by

$$\tilde{f}_g(t,\omega) = \mathscr{F}\{f_t(\tau)\} = \hat{f}_t(\upsilon), \tag{4.3.4}$$

where $f_t(\tau) = f(\tau)\,g(\tau - t)$.

2. If the window g is real and symmetric with $g(\tau) = g(-\tau)$ and if g is normalized so that $\|g\| = 1$ and $\|g_{t,\omega}\| = \|g(\tau - t)\| = 1$ for any $(t,\omega) \in \mathbb{R}^2$, then the Gabor transform of $f \in L^2(\mathbb{R})$ becomes

$$\tilde{f}_g(t,\omega) = \langle f, g_{t,\omega}\rangle = \int_{-\infty}^{\infty} f(\tau)\,g(\tau - t)\,e^{-i\omega\tau}d\tau. \tag{4.3.5}$$

This can be interpreted as the *short-time Fourier transform* because the multiplication by $g(\tau - t)$ induces localization of the Fourier integral in the neighborhood of $\tau = t$. Application of the Schwarz inequality (2.6.1) to (4.3.5) gives

$$\left|\tilde{f}_g(t,\omega)\right| = |\langle f, g_{t,\omega}\rangle| \le \|f\|\,\|g_{t,\omega}\| = \|f\|\,\|g\|.$$

This shows that the Gabor transform $\tilde{f}_g(t,\omega)$ is *bounded*.

3. The energy density defined by

$$\left|\tilde{f}_g(t,\omega)\right|^2 = \left|\int_{-\infty}^{\infty} f(\tau)\,g(\tau - t)\,e^{-i\omega\tau}d\tau\right|^2 \tag{4.3.6}$$

measures the energy of a signal in the time–frequency plane in the neighborhood of the point (t,ω).

4. It follows from definition (4.3.3) with a fixed ω that

$$\tilde{f}_g(t,\omega) = e^{-i\omega t}\int_{-\infty}^{\infty} f(\tau)\,g(\tau - t)\,e^{i\omega(t-\tau)}d\tau = e^{-i\omega t}\big(f * g_\omega\big)(t), \tag{4.3.7}$$

where $g_\omega(\tau) = e^{i\omega\tau}g(\tau)$ and $g(-\tau) = g(\tau)$. Furthermore, by the Parseval relation (3.4.34) of the Fourier transform, we find

$$\tilde{f}_g(t,\omega) = \langle f, \overline{g}_{t,\omega}\rangle = \left\langle \hat{f}, \hat{\overline{g}}_{t,\omega}\right\rangle = e^{i\omega t}\int_{-\infty}^{\infty} \hat{f}(\upsilon)\,\hat{g}(\upsilon - \omega)\,e^{-i\upsilon t}d\upsilon. \tag{4.3.8}$$

Except for the factor $\exp(i\omega t)$, result (4.3.8) is almost identical with (4.3.3), but the time variable t is replaced by the frequency variable ω, and the time window $g(\tau - t)$ is replaced by the frequency window $\hat{g}(\upsilon - \omega)$. The extra factor $\exp(i\omega t)$ in (4.3.8) is associated with the Weyl commutation relations of the Weyl–Heisenberg group which describe translations in time and frequency. If the window is well localized in frequency and in time, that is, if $\hat{g}(\upsilon - \omega)$ is small outside a small frequency band in addition to $g(\tau)$ being small outside a small time interval, then (4.3.8) reveals that the Gabor transform gives a local time–frequency analysis of the signal f in the sense that it provides accurate information of f simultaneously in both time and frequency domains. However, all functions, including the window function, satisfy the Heisenberg uncertainty principle, that is, the joint resolution $\sigma_t \sigma_\omega$ of a signal cannot be arbitrarily small and has always greater than the minimum value $\frac{1}{2}$ which is attained only for the Gaussian window function $g(t) = \exp\left(-at^2\right)$.

5. For a fixed ω, the Fourier transform of $\tilde{f}_g(t,\omega)$ with respect to t is given by the following:

$$\mathscr{F}\left\{\tilde{f}_g(t,\omega)\right\} = \hat{\tilde{f}}_g(\upsilon,\omega) = \hat{f}(\upsilon+\omega)\,\hat{g}(\upsilon). \tag{4.3.9}$$

This follows from the Fourier transform of (4.3.7) with respect to t

$$\mathscr{F}\left\{\tilde{f}_g(t,\omega)\right\} = \mathscr{F}\left\{e^{-i\omega t}\left(f * g_\omega\right)(t)\right\} = \hat{f}(\upsilon+\omega)\,\hat{g}(\upsilon).$$

6. If $g(t) = \exp\left(-\frac{1}{4}t^2\right)$, then

$$\tilde{f}_g(t,\omega) = \sqrt{2}\,\exp\left(i\omega t - \omega^2\right)\left(Wf\right)(t+2i\omega), \tag{4.3.10}$$

where W represents the *Weierstrass transformation* of $f(x)$ defined by

$$W\left[f(x)\right] = \frac{1}{2\sqrt{2}}\int_{-\infty}^{\infty} f(x)\exp\left[-\frac{1}{4}(t-x)^2\right]dx. \tag{4.3.11}$$

7. The time width σ_t around t and the frequency spread σ_ω around ω are independent of t and ω. We have, by definition, and the Gabor window function (4.3.1),

$$\sigma_t^2 = \int_{-\infty}^{\infty}(\tau-t)^2\left|g_{t,\omega}(\tau)\right|^2 d\tau = \int_{-\infty}^{\infty}(\tau-t)^2\left|g(\tau-t)\right|^2 d\tau = \int_{-\infty}^{\infty}\tau^2\left|g(\tau)\right|^2 d\tau.$$

Similarly, we obtain, by (4.3.2),

$$\sigma_\omega^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty}(\upsilon-\omega)^2\left|\hat{g}_{t,\omega}(\upsilon)\right|^2 d\upsilon = \frac{1}{2\pi}\int_{-\infty}^{\infty}(\upsilon-\omega)^2\left|\hat{g}(\upsilon)\right|^2 d\upsilon = \frac{1}{2\pi}\int_{-\infty}^{\infty}\upsilon^2\left|\hat{g}(\upsilon)\right|^2 d\upsilon.$$

Thus, both σ_t and σ_ω are independent of t and ω. The energy spread of $g_{t,\omega}(\tau)$ can be represented by the Heisenberg rectangle centered at (t,ω) with the area $\sigma_t\sigma_\omega$ which is independent of t and ω. This means that the Gabor transform has the same resolution in the time–frequency plane.

Example 4.3.1. Obtain the Gabor transform of functions
(a) $f(\tau) = 1$, (b) $f(\tau) = \exp(-i\sigma\tau)$.

We obtain

$$\text{(a)} \quad \tilde{f}_g(t,\omega) = \int_{-\infty}^{\infty} g(\tau - t)\, e^{-i\omega\tau} d\tau = e^{-i\omega t}\hat{g}(\omega).$$

$$\text{(b)} \quad \tilde{f}_g(t,\omega) = \int_{-\infty}^{\infty} e^{-i\tau(\omega+\sigma)} g(\tau - t)\, d\tau = \exp\big\{-it(\sigma+\omega)\big\}\hat{g}(\sigma+\omega).$$

Example 4.3.2. Find the Gabor transform of functions
(a) $f(\tau) = \delta(\tau)$, (b) $f(\tau) = \delta(\tau - t_0)$.
We have

$$\text{(a)} \quad \tilde{f}_g(t,\omega) = \int_{-\infty}^{\infty} \delta(\tau)\, g(\tau - t)\, e^{-i\omega\tau} d\tau = g(-t).$$

$$\text{(b)} \quad \tilde{f}_g(t,\omega) = \int_{-\infty}^{\infty} \delta(\tau - t_0)\, g(\tau - t)\, e^{-i\omega\tau} d\tau = e^{-i\omega t_0} g(t_0 - t).$$

Example 4.3.3. Find the Gabor transform of the function $f(\tau) = \exp\left(-a^2\tau^2\right)$ with $g(\tau) = 1$.
We have

$$\tilde{f}_g(t,\omega) = \int_{-\infty}^{\infty} \exp\big\{-\big(a^2\tau^2 + i\omega\tau\big)\big\}\, d\tau = \hat{f}(\omega) = \frac{\sqrt{\pi}}{a}\exp\left(-\frac{\omega^2}{4a^2}\right).$$

4.4 Basic Properties of Gabor Transforms

Theorem 4.4.1 (Linearity). *If the Gabor transforms of two functions f_1 and f_2 exist with respect to a window function g, then*

$$\mathcal{G}\big[af_1 + bf_2\big](t,\omega) = a\,\mathcal{G}\big[f_1\big](t,\omega) + b\,\mathcal{G}\big[f_2\big](t,\omega), \tag{4.4.1}$$

where a and b are two arbitrary constants.

The proof easily follows from the definition of the Gabor transform and is left as an exercise.

Theorem 4.4.2. *If f and $g \in L^2(\mathbb{R})$, then the following results hold:*

$$(a) \quad (\textit{Translation}) : \mathcal{G}\big[T_a f\big](t,\omega) = e^{-i\omega a}\,\mathcal{G}\big[f\big](t-a,\omega), \tag{4.4.2}$$

$$(b) \quad (\textit{Modulation}) : \mathcal{G}\big[M_a f\big](t,\omega) = \mathcal{G}\big[f\big](t,\omega - a), \tag{4.4.3}$$

$$(c) \quad (\textit{Conjugation}) : \mathcal{G}\big[\bar{f}\,\big](t,\omega) = \overline{\mathcal{G}}\big[f\big](\tau,-\omega). \tag{4.4.4}$$

Proof. (a) We have, by definition,

$$\mathcal{G}\big[T_a f\big](t,\omega) = \mathcal{G}\big[f(\tau - a)\big](t,\omega) = \int_{-\infty}^{\infty} f(\tau - a)\, g(\tau - t)\, e^{-i\omega\tau} d\tau$$

$$= e^{-i\omega a} \int_{-\infty}^{\infty} f(x)\, g\big(x - \overline{t-a}\big)\, e^{-i\omega x} dx$$
$$= e^{-i\omega a} \mathscr{G}\big[f\big](t-a, \omega).$$

(b) We have

$$\begin{aligned}\mathscr{G}\big[M_a f\big](t,\omega) &= \mathscr{G}\big[e^{ia\tau} f(\tau)\big](t,\omega)\\ &= \int_{-\infty}^{\infty} f(\tau)\, g(\tau - t)\, e^{-i\tau(\omega - a)} d\tau\\ &= \mathscr{G}\big[f\big](t, \omega - a).\end{aligned}$$

(c) It follows from definition (4.3.3) with a real window function g that

$$\mathscr{G}\big[\bar{f}\,\big](t,\omega) = \int_{-\infty}^{\infty} \overline{f(\tau)}\, g(\tau - t)\, e^{-i\omega\tau} d\tau = \overline{\int_{-\infty}^{\infty} f(\tau)\, g(\tau - t)\, e^{i\omega\tau} d\tau}$$
$$= \overline{\mathscr{G}\big[f\big](\tau, -\omega)}.$$

Theorem 4.4.3. *If two signals $f, g \in L^2(\mathbb{R})$, then*

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left|\tilde{f}_g(t,\omega)\right|^2 dt\, d\omega = \|f\|_2^2 \|g\|_2^2.$$

Proof. The left-hand side of the above result is equal to

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left|\tilde{f}_g(t,\omega)\right|^2 dt\, d\omega = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left|\int_{-\infty}^{\infty} f(\tau)\, g(\tau - t)\, e^{-i\omega\tau} d\tau\right|^2 dt\, d\omega$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left|\int_{-\infty}^{\infty} h_t(\tau)\, e^{-i\omega\tau} d\tau\right|^2 dt\, d\omega, \quad h_t(\tau) = f(\tau)\, g(\tau - t)$$
$$= \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} \left|\hat{h}_t(\omega)\right|^2 d\omega$$
$$= \int_{-\infty}^{\infty} \left\|\hat{h}_t(\omega)\right\|^2 dt$$
$$= \int_{-\infty}^{\infty} \left\|h_t(\tau)\right\|^2 dt, \qquad \text{by Plancherel's theorem}$$
$$= \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} \left|f(\tau)\right|^2 \left|g(\tau - t)\right|^2 d\tau$$

$$= \int_{-\infty}^{\infty} |f(\tau)|^2 d\tau \int_{-\infty}^{\infty} |g(x)|^2\, dx = \|f\|_2^2 \|g\|_2^2.$$

This completes the proof.

Theorem 4.4.4 (Parseval's Formula). *If* $\mathscr{G}[f](t,\omega) = \tilde{f}_g(t,\omega)$ *and* $\mathscr{G}[h](t,\omega) = \tilde{h}_g(t,\omega)$, *then the Parseval formula for the Gabor transform is given by*

$$\langle \tilde{f}, \tilde{h} \rangle = \|g\|^2 \langle f, h \rangle, \tag{4.4.5}$$

where

$$\langle \tilde{f}, \tilde{h} \rangle = \langle \tilde{f}, \tilde{h} \rangle_{L^2(\mathbb{R}^2)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{f}_g(t,\omega)\, \overline{\tilde{h}}_g(t,\omega)\, dt\, d\omega. \tag{4.4.6}$$

In particular, if $\|g\| = 1$, *then the Gabor transformation is an isometry from* $L^2(\mathbb{R})$ *into* $L^2(\mathbb{R}^2)$.

Proof. We first note that, for a fixed t,

$$\tilde{f}_g(t,\omega) = \mathscr{F}\{f_t(\tau)\} = \mathscr{F}\{f(\tau) g_t(\tau)\},$$

where $g_t(\tau) = g(\tau - t)$.

Thus, the Parseval formula (3.4.34) for the Fourier transform gives

$$\begin{aligned}
\int_{-\infty}^{\infty} \tilde{f}(t,\omega)\, \overline{\tilde{h}}(t,\omega)\, d\omega &= \left\langle \mathscr{F}\{f g_t\}, \mathscr{F}\{h g_t\} \right\rangle \\
&= \langle f g_t, h g_t \rangle = \int_{-\infty}^{\infty} f(\tau)\, g(\tau - t)\, \overline{h}(\tau)\, \overline{g(\tau - t)}\, d\tau \\
&= \int_{-\infty}^{\infty} f(\tau)\, \overline{h}(\tau)\, |g(\tau - t)|^2\, d\tau.
\end{aligned}$$

Integrating this result with respect to t from $-\infty$ to ∞ gives

$$\begin{aligned}
\langle \tilde{f}, \tilde{h} \rangle &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{f}(t,\omega)\, \overline{\tilde{h}}(t,\omega)\, dt\, d\omega \\
&= \int_{-\infty}^{\infty} f(\tau)\, \overline{h}(\tau)\, d\tau \int_{-\infty}^{\infty} |g(\tau - t)|^2\, dt \\
&= \int_{-\infty}^{\infty} f(\tau)\, \overline{h}(\tau)\, d\tau \int_{-\infty}^{\infty} |g(x)|^2\, dx \qquad (\tau - t = x) \\
&= \|g\|^2 \langle f, h \rangle.
\end{aligned}$$

This proves the result.

If $\|g\| = 1$, then (4.4.5) shows isometry from $L^2(\mathbb{R})$ into $L^2(\mathbb{R}^2)$.

Theorem 4.4.5 (Inversion Theorem). *If a function $f \in L^2(\mathbb{R})$, then*

$$f(\tau) = \frac{1}{2\pi} \frac{1}{\|g\|^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{f}_g(t,\omega)\, g(\tau - t)\, e^{i\omega\tau}\, d\omega\, dt. \tag{4.4.7}$$

First Proof. It follows from the continuous Gabor transform (4.3.3) that

$$\tilde{f}_g(t,\omega) = \mathscr{F}\{f(\tau)g(\tau - t)\},$$

where the Fourier transform with respect to τ is taken.

Application of the inverse Fourier transform to this result gives

$$f(\tau)\, g(\tau - t) = \mathscr{F}^{-1}\left\{\tilde{f}_g(t,\omega)\right\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega\tau} \tilde{f}_g(t,\omega)\, d\omega.$$

Multiplying this result by $\overline{g}(\tau - t)$ and integrating with respect to t yields

$$f(\tau) \int_{-\infty}^{\infty} \left|g(\tau - t)\right|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i\omega\tau}\, \overline{g}(\tau - t)\, \tilde{f}_g(t,\omega)\, d\omega\, dt.$$

Or, equivalently,

$$f(\tau)\|g\|^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i\omega\tau}\, \overline{g}(\tau - t)\, \tilde{f}_g(t,\omega)\, d\omega\, dt.$$

This proves the inversion theorem.

Second Proof. We apply the inverse Fourier transform of $f(\tau)$ and the Parseval formula to replace $\|g\|^2$ by $\frac{1}{2\pi}\|\hat{g}\|^2$ so that

$$\begin{aligned} f(\tau)\|g\|^2 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega\tau}\, \hat{f}(\omega)\, d\omega \cdot \frac{1}{2\pi}\|\hat{g}\|^2 \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega\tau}\, \hat{f}(\omega)\, d\omega \cdot \frac{1}{2\pi} \int_{-\infty}^{\infty} \left|\hat{g}(\upsilon)\right|^2 d\upsilon. \end{aligned}$$

Since the integral is true for any arbitrary ω, we replace ω by ω + υ to obtain

$$\begin{aligned}
f(\tau)\|g\|^2 &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\tau(\omega+\upsilon)}\,\hat{f}(\omega+\upsilon)d\omega \cdot \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{g}(\upsilon)\hat{g}(\upsilon)d\upsilon \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega\tau}d\omega \cdot \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\tau\upsilon}\big[\hat{f}(\omega+\upsilon)\,\hat{g}(\upsilon)\big]\hat{g}(\upsilon)\,d\upsilon \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega\tau}d\omega \cdot \left[\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\tau\upsilon}\,\hat{\tilde{f}}_g(\omega+\upsilon)\,\hat{g}(\upsilon)\,d\upsilon\right], \quad \text{by (4.3.9)} \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega\tau}d\omega \cdot \left[\tilde{f}_g(\tau,\omega) * g(\tau)\right], \quad \text{by (3.3.23)} \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega\tau}d\omega \int_{-\infty}^{\infty} \tilde{f}_g(t,\omega)\,g(\tau-t)\,dt \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i\omega\tau}\,\tilde{f}_g(t,\omega)\,g(\tau-t)\,dt\,d\omega.
\end{aligned}$$

This proves the inversion theorem.

Theorem 4.4.6 (Conservation of Energy). *If* $f \in L^2(\mathbb{R})$, *then*

$$\|f\|_2^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left|\tilde{f}_g(t,\omega)\right|^2 dt\,d\omega, \tag{4.4.8}$$

where g *is a normalized window function* $\left(\|g\| = 1\right)$.

Proof. Using (4.3.9) dealing with the Fourier transform of $\tilde{f}_g(t,\omega)$ with respect to t, we apply the Plancherel formula to the right-hand side of (4.4.8) to obtain

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left|\tilde{f}_g(t,\omega)\right|^2 dt\,d\omega &= \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\,\frac{1}{2\pi}\int_{-\infty}^{\infty} \left|\mathcal{F}\left\{\tilde{f}_g(t,\omega)\right\}\right|^2 d\upsilon \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\,\frac{1}{2\pi}\int_{-\infty}^{\infty} \left|\hat{f}(\omega+\upsilon)\right|^2 |\hat{g}(\upsilon)|^2\,d\upsilon \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\,\frac{1}{2\pi}\int_{-\infty}^{\infty} \left|\hat{f}(\omega)\right|^2 |\hat{g}(\upsilon)|^2\,d\upsilon \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \left|\hat{f}(\omega)\right|^2 d\omega, \quad \text{since } \|\hat{g}\| = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\hat{g}(\upsilon)|^2\,d\upsilon = 1 \\
&= \int_{-\infty}^{\infty} |f(\tau)|^2\,d\tau = \|\hat{f}\|_2^2.
\end{aligned}$$

This completes the proof.

Physically, the Gabor transformation transforms a signal f of one variable τ to a function $\tilde{f}$ of two variables t and ω *without* changing its total energy.

4.5 Frames and Frame Operators

The concept of frames in a Hilbert space was originally introduced by Duffin and Schaeffer (1952) in the context of nonharmonic Fourier series only 6 years after Gabor (1946) published his famous work. In signal processing, this concept has become useful in analyzing the completeness and stability of linear discrete signal representations. A frame is a set of vectors $\{\phi_n\}_{n\in\Gamma}$ that characterizes any signal f from its inner products $\{\langle f, \phi_n\rangle\}_{n\in\Gamma}$, where Γ is the index set, which may be finite or infinite.

Definition 4.5.1 (Basis). A sequence of vectors $\{x_n\}$ in a Hilbert space H is called a *basis* (*Schauder basis*) of H if to each $x \in H$, there corresponds a unique sequence of scalars $\{a_n\}_{n=1}^{\infty}$ such that

$$x = \sum_{n=1}^{\infty} a_n x_n, \tag{4.5.1}$$

where the convergence is defined by the norm.

Definition 4.5.2 (Orthogonal Basis and Orthonormal Basis). A basis $\{x_n\}_{n=1}^{\infty}$ of H is called *orthogonal* if $\langle x_n, x_m\rangle = 0$ for $n \neq m$.

An orthogonal basis is called *orthonormal* if $\langle x_n, x_n\rangle = 1$ for all n.

An orthogonal basis $\{x_n\}_{n=1}^{\infty}$ is complete in the sense that if $\langle x, x_n\rangle = 0$ for all n, then $x = 0$ (see Theorem 2.10.4).

Every separable Hilbert space has an orthonormal basis, and for an orthonormal basis the expansion (4.5.1) has the form

$$x = \sum_{n=1}^{\infty} \langle x, x_n\rangle x_n, \tag{4.5.2}$$

with

$$\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, x_n\rangle|^2. \tag{4.5.3}$$

More generally, for any $x, y \in H$,

$$\langle x, y\rangle = \sum_{n=1}^{\infty} \langle x, x_n\rangle \overline{\langle y, x_n\rangle}. \tag{4.5.4}$$

It can be proved that every basis $\{x_n\}_{n=1}^{\infty}$ of a Hilbert space H possesses a unique *biorthogonal basis* $\{x_n^*\}_{n=1}^{\infty}$ which implies that

$$\langle x_m, x_n^*\rangle = \delta_{m,n},$$

and, for every $x \in H$, we have

$$x = \sum_{n=1}^{\infty} \langle x, x_n^* \rangle x_n = \sum_{n=1}^{\infty} \langle x, x_n \rangle x_n^*.$$

If $\langle x_m, x_n^* \rangle x_n = 0$ for $m \neq n$, but $\langle x_n, x_n^* \rangle x_n$ is not necessarily equal to one, $\{x_n^*\}_{n=1}^{\infty}$ is called a *biorthogonal basis* of $\{x_n\}_{n=1}^{\infty}$. In this case, we have, for any $x \in H$,

$$x = \sum_{n=1}^{\infty} (\overline{a}_n)^{-1} \langle x, x_n^* \rangle x_n = \sum_{n=1}^{\infty} (\overline{a}_n)^{-1} \langle x, x_n \rangle x_n^*, \tag{4.5.5}$$

where $a_n = \langle x_n^*, x_n \rangle \neq 0$. It can be shown that $a_n \neq 0$ for all n.

Definition 4.5.3 (Bounded Basis, Unconditional Basis, and Riesz Basis). If $\{x_n\}$ is a basis in a separable Hilbert space H, then

(i) $\{x_n\}$ is called a *bounded basis* if there exist two nonnegative numbers A and B such that

$$A \leq \|x_n\| \leq B \qquad \text{for all } n.$$

(ii) $\{x_n\}$ is called an *unconditional basis* in a separable Hilbert space H if

$$\sum a_n x_n \in H \quad \text{implies that} \quad \sum |a_n| x_n \in H.$$

(iii) $\{x_n\}$ is called a *Riesz basis* if there exist a topological isomorphism $T : H \to H$ and an orthonormal basis $\{y_n\}$ of H such that $Tx_n = y_n$ for every n.

Remark. In a Hilbert space, all bounded unconditional bases are equivalent to an orthonormal basis. In other words, if $\{x_n\}$ is a bounded unconditional basis, then there exists an orthonormal basis $\{e_n\}$ and a topological isomorphism $T : H \to H$ such that $Te_n = y_n$ for all n.

Definition 4.5.4 (Frame). A sequence $\{x_n\}$ in a separable Hilbert space H (not necessarily a basis of H) is called a *frame* if there exist two numbers A and B with $0 < A \leq B < \infty$ such that

$$A\|x\|^2 \leq \sum_n |\langle x, x_n \rangle|^2 \leq B\|x\|^2. \tag{4.5.6}$$

The numbers A and B are called the *frame bounds*. If $A = B$, the frame is said to be *tight*. The frame is called *exact* if it ceases to be a frame whenever any single element is deleted from the frame.

Definition 4.5.5 (Frame Operator). To each frame $\{x_n\}$ there corresponds an operator T, called the *frame operator*, from H into itself defined by

$$Tx = \sum_n \langle x, x_n \rangle x_n \qquad \text{for all } x \in H. \tag{4.5.7}$$

Remark. The x_n's are not necessarily linearly independent. Since $\sum_n |\langle x, x_n\rangle|^2$ is a series of positive real numbers, it converges absolutely and hence, unconditionally.

The following example shows that tightness and exactness are not related.

Example 4.5.1. If $\{e_n\}_{n=1}^{\infty}$ is an orthonormal basis of H, then

(i) $\{e_1, e_1, e_2, e_2, \ldots\}$ is a tight frame with frame bounds $A = 2 = B$, but it is not exact.

(ii) $\left\{\sqrt{2}e_1, e_2, e_3, \ldots\right\}$ is an exact frame but not tight since the frame bounds are easily seen as $A = 1$ and $B = 2$.

(iii) $\left\{e_1, \frac{e_2}{\sqrt{2}}, \frac{e_2}{\sqrt{2}}, \frac{e_3}{\sqrt{3}}, \frac{e_3}{\sqrt{3}}, \frac{e_3}{\sqrt{3}}, \ldots\right\}$ is a tight frame with the frame bound $A = 1$ but not an orthonormal basis.

(iv) $\left\{e_1, \frac{e_2}{2}, \frac{e_3}{3}, \ldots\right\}$ is a complete orthogonal sequence but is not a frame.

If $\{x_n\}$ is an orthonormal basis of H, then the Parseval formula holds, that is, for any $x \in H$,

$$\|x\|^2 = \sum_n |\langle x, x_n\rangle|^2.$$

It follows from the definition of frame that $\{x_n\}$ is a tight frame with frame bounds $A = B = 1$.

But the converse is not necessarily true. That is, tight frames are not necessarily orthonormal. For example, $H = \mathbb{C}^2$ and

$$e_1 = (0, 1), e_2 = \left(\frac{\sqrt{3}}{2}, -\frac{1}{2}\right), e_3 = \left(-\frac{\sqrt{3}}{2}, -\frac{1}{2}\right).$$

For any $x = (x_1, x_2) \in H$, we have

$$\begin{aligned}\sum_{i=1}^{3} |\langle x, e_i\rangle|^2 &= |x_2|^2 + \left|\frac{\sqrt{3}}{2}x_1 - \frac{1}{2}x_2\right|^2 + \left|-\frac{\sqrt{3}}{2}x_1 - \frac{1}{2}x_2\right|^2 \\ &= |x_2|^2 + \frac{1}{2}\left(3|x_1|^2 + |x_2|^2\right) \\ &= \frac{3}{2}\left(|x_1|^2 + |x_2|^2\right) = \frac{3}{2}\|x\|^2.\end{aligned}$$

Thus, three vectors (e_1, e_2, e_3) define a tight frame with the frame bounds $A = B = \frac{3}{2}$ but they are not orthonormal since (e_1, e_2, e_3) are not linearly independent.

Theorem 4.5.1. *If a sequence $\{x_n\}$ is a tight frame in H with the frame bound $A = 1$, and if $\|x_n\| = 1$ for all n, then $\{x_n\}$ is an orthonormal basis of H.*

Proof. It follows from (4.5.6) that

$$\|x_m\|^2 = \sum_n |\langle x_m, x_n\rangle|^2 = \|x_m\|^4 + \sum_{m\neq n} |\langle x_m, x_n\rangle|^2.$$

Since $\|x_m\| = 1$, the above equality implies that

$$\langle x_m, x_n\rangle = 0 \qquad \text{for } m \neq n.$$

The completeness of $\{x_n\}$ is a consequence of the fact that frames are complete.

To check this, suppose $x \in H$ such that $\langle x, x_n\rangle = 0$ for all n. Then, the relation

$$A\|x\|^2 \leq \sum_n |\langle x, x_n\rangle|^2 = 0$$

implies that $x = 0$.

Theorem 4.5.2. *Suppose a sequence $\{x_n\}$ is a separable Hilbert space H. Then, the following are equivalent.*

(a) The frame operator $Tx = \sum_n \langle x, x_n\rangle x_n$ is a bounded linear operator on it with $AI \leq T \leq BI$, where I is the identity operator on H.

(b) $\{x_n\}_{n=1}^{\infty}$ is a frame with frame bounds A and B.

Proof. If (a) holds, then the relation $AI \leq T \leq BI$ is equivalent to

$$\langle AIx, x\rangle \leq \langle Tx, x\rangle \leq \langle BIx, x\rangle \qquad \text{for all } x \in H. \tag{4.5.8}$$

Since I is an identity operator, $\langle Ix, x\rangle = \|x\|^2$. Also,

$$\begin{aligned}\langle Tx, x\rangle &= \left\langle \sum_n \langle x, x_n\rangle x_n, x \right\rangle = \sum_n \langle x, x_n\rangle\langle x_n, x\rangle \\ &= \sum_n \langle x, x_n\rangle\overline{\langle x, x_n\rangle} = \sum_n |\langle x, x_n\rangle|^2.\end{aligned}$$

Evidently, inequality (4.5.8) gives

$$A\|x\|^2 \leq \sum_n |\langle x, x_n\rangle|^2 \leq B\|x\|^2.$$

This shows that (a) implies (b).

We next prove that (b) implies (a). Suppose (b) holds, that is, $\{x_n\}$ is a frame with frame bounds A and B. Recall that in any Hilbert space H the norm of any element $x \in H$ is given by

$$\|x\| = \sup_{\|y\|=1} |\langle x, y\rangle| \qquad \text{for } y \in H.$$

For a fixed $x \in H$, we consider

$$T_N x = \sum_{n=-N}^{N} \langle x, x_n \rangle x_n.$$

For $0 \leq M \leq N$, we have, by the Schwarz inequality,

$$\begin{aligned}
\left\| T_N x - T_M x \right\|^2 &= \sup_{\|y\|=1} \left| \langle T_N x - T_M x, y \rangle \right|^2 \\
&= \sup_{\|y\|=1} \left| \sum_{M+1 \leq |n| \leq N} \langle x, x_n \rangle \langle x_n, y \rangle \right|^2 \\
&\leq \sup_{\|y\|=1} \left(\sum_{M+1 \leq |n| \leq N} \left| \langle x, x_n \rangle \right|^2 \right) \left(\sum_{M+1 \leq |n| \leq N} \left| \langle x_n, y \rangle \right|^2 \right) \\
&\leq \sup_{\|y\|=1} \left(\sum_{M+1 \leq |n| \leq N} \left| \langle x, x_n \rangle \right|^2 \right) B \left\| y \right\|^2, \quad \text{by (4.5.6)} \\
&= B \sup_{\|y\|=1} \left(\sum_{M+1 \leq |n| \leq N} \left| \langle x, x_n \rangle \right|^2 \right) \to 0 \quad \text{as } M, N \to \infty.
\end{aligned}$$

Thus, $\{T_N x\}$ is a Cauchy sequence in H and hence it is convergent as $N \to \infty$. Therefore,

$$\lim_{N \to \infty} T_N x = TX.$$

Next, we use the preceding argument to obtain

$$\begin{aligned}
\left\| Tx \right\|^2 = \sup_{\|y\|=1} \left| \langle Tx, y \rangle \right|^2 &= \sup_{\|y\|=1} \left| \left\langle \sum_n \langle x, x_n \rangle x_n, x \right\rangle \right|^2 \\
&= \sup_{\|y\|=1} \left| \sum_n \langle x, x_n \rangle \langle x_n, x \rangle \right|^2 \\
&\leq B \left(\sum_n \left| \langle x, x_n \rangle \right|^2 \right) \leq B^2 \left\| x \right\|^2.
\end{aligned}$$

This implies that $\|T\| \leq B$, and hence the frame operator T is bounded.

Since $\langle Ix, x \rangle = \|x\|^2$, it follows from definition (4.5.6) that

$$A \langle Ix, x \rangle \leq \langle Tx, x \rangle \leq B \langle Ix, x \rangle,$$

which is equivalent to the relation

$$AI \le T \le BI.$$

This complete the proof.

Theorem 4.5.3. *Suppose* $\{x_n\}_{n=1}^{\infty}$ *is a frame on a separable Hilbert space with frame bounds* A *and* B*, and* T *is the corresponding frame operator. Then,*

(a) T *is invertible and* $B^{-1}I \le T^{-1} \le A^{-1}I$*. Furthermore,* T^{-1} *is a positive operator and hence it is self-adjoint.*
(b) $\{T^{-1}x_n\}$ *is a frame with frame bounds* B^{-1} *and* A^{-1} *with* $A^{-1} \ge B^{-1} > 0$*, and it is called the dual frame of* $\{x_n\}$*.*
(c) Every $x \in H$ *can be expressed in the form*

$$x = \sum_n \langle x, T^{-1}x_n\rangle x_n = \sum_n \langle x, x_n\rangle T^{-1}x_n. \tag{4.5.9}$$

The frame $\{T^{-1}x_n\} = \{\tilde{x}_n\}$ *is called the dual frame of* $\{x_n\}$*. It is easy to verify that the dual frame of* $\{\tilde{x}_n\}$ *is the original frame* $\{x_n\}$*. According to formula (4.5.9), the reconstruction formula for* x *has the form*

$$x = \sum_n \langle x, \tilde{x}_n\rangle x_n = \sum_n \langle x, x_n\rangle \tilde{x}_n. \tag{4.5.10}$$

Proof. (a) Since the frame operator T satisfies the relation

$$AI \le T \le BI, \tag{4.5.11}$$

it follows that

$$\left(I - B^{-1}T\right) \le \left(I - B^{-1}AI\right) = \left(I - \frac{A}{B}\right) I$$

and hence

$$\left\|I - B^{-1}T\right\| \le \left\|\left(I - \frac{A}{B}\right)\right\| < 1.$$

Thus, $B^{-1}T$ is invertible and consequently so is T. We next multiply (4.5.11) by T^{-1} and use the fact that T^{-1} commutes with I and T to obtain

$$B^{-1}I \le T \le A^{-1}I.$$

In view of the fact that

$$\left\langle T^{-1}x, x\right\rangle = \left\langle T^{-1}x, T(T^{-1}x)\right\rangle \ge A\left\langle T^{-1}x, T^{-1}x\right\rangle = A\left\|T^{-1}x\right\|^2 \ge 0,$$

we conclude that T^{-1} is a positive operator and hence it is self-adjoint.

(b) Since T^{-1} is self-adjoint, we have

$$\sum_n \langle x, T^{-1}x_n\rangle T^{-1}x_n = T^{-1}\left(\sum_n \langle T^{-1}x, x_n\rangle x_n\right) = T^{-1}\Big(T\big(T^{-1}x\big)\Big) = T^{-1}x. \tag{4.5.12}$$

This gives

$$\left\langle \sum_n \langle x, T^{-1}x_n\rangle T^{-1}x_n, x\right\rangle = \langle T^{-1}x, x\rangle.$$

Or,

$$\sum_n \langle x, T^{-1}x_n\rangle\langle T^{-1}x_n, x\rangle = \langle T^{-1}x, x\rangle.$$

Hence,

$$\sum_n \langle x, T^{-1}x_n\rangle\overline{\langle x, T^{-1}x_n\rangle} = \langle T^{-1}x, x\rangle.$$

Or,

$$\sum_n \left|\langle x, T^{-1}x_n\rangle\right|^2 = \langle T^{-1}x, x\rangle.$$

Using the result from (a), that is, $B^{-1}I \le T \le A^{-1}I$, it turns out that

$$B^{-1}\langle Ix, x\rangle \le \langle T^{-1}x, x\rangle \le A^{-1}\langle Ix, x\rangle$$

and hence

$$B^{-1}\|x\|^2 \le \sum_n \left|\langle x, T^{-1}x_n\rangle\right|^2 \le A^{-1}\|x\|^2. \tag{4.5.13}$$

This shows that $\{T^{-1}x_n\}$ is a frame with frame bounds B^{-1} and A^{-1}.

(c) We replace x by $T^{-1}x$ in (4.5.7) to derive

$$x = \sum_n \langle T^{-1}x, x_n\rangle x_n = \sum_n \langle x, T^{-1}x_n\rangle x_n.$$

Similarly, replacing x by Tx in (4.5.12) gives

$$x = \sum_n \langle Tx, T^{-1}x_n\rangle T^{-1}x_n = \sum_n \langle T^{-1}Tx, x_n\rangle T^{-1}x_n = \sum_n \langle x, x_n\rangle T^{-1}x_n.$$

This completes the proof.

Theorem 4.5.4. *Suppose* $\{x_n\}_{n=1}^{\infty}$ *is a frame on a separable Hilbert space* H *with frame bounds* A *and* B*. If there exists a sequence of scalars* $\{c_n\}$ *such that* $x = \sum_n c_n x_n$*, then*

$$\sum_n |c_n|^2 = \sum_n |a_n|^2 + \sum_n |a_n - c_n|^2, \tag{4.5.14}$$

where $a_n = \langle x, T^{-1}x_n \rangle$ *so that* $x = \sum_n a_n x_n$.

Proof. Note that $\langle x_n, T^{-1}x \rangle = \langle T^{-1}x_n, x \rangle = \overline{a}_n$. Substituting $x = \sum_n a_n x_n$ into the first term in the inner product $\langle x, T^{-1}x \rangle$ gives

$$\langle x, T^{-1}x \rangle = \left\langle \sum_n a_n x_n, T^{-1}x \right\rangle = \sum_n a_n \langle x_n, T^{-1}x \rangle = \sum_n |a_n|^2.$$

Similarly, substituting $x = \sum_n c_n x_n$ into the first term in $\langle x, T^{-1}x \rangle$ yields

$$\langle x, T^{-1}x \rangle = \left\langle \sum_n c_n x_n, T^{-1}x \right\rangle = \sum_n c_n \langle x_n, T^{-1}x \rangle = \sum_n c_n \overline{a}_n.$$

Consequently,

$$\sum_n |a_n|^2 = \sum_n c_n \overline{a}_n. \tag{4.5.15}$$

Finally, we obtain, by using (4.5.15),

$$\sum_n |a_n|^2 + \sum_n |a_n - c_n|^2 = \sum_n |a_n|^2 + \sum_n \left(|a_n|^2 - a_n \overline{c}_n - \overline{a}_n c_n + |c_n|^2 \right) = \sum_n |c_n|^2.$$

This completes the proof.

Theorem 4.5.5. *A necessary and sufficient condition for a sequence* $\{x_n\}$ *on a Hilbert space* H *to be an exact frame is that the sequence* $\{x_n\}$ *be a bounded unconditional basis of* H.

Proof. The condition is necessary.

We assume that $\{x_n\}$ is an exact frame with frame bounds A and B. Then, $\{x_n\}$ and $\{T^{-1}x_n\}$ are biorthonormal. For a fixed m, we have

$$A \|T^{-1}x_m\|^2 \le \sum_n |\langle T^{-1}x_m, x_n \rangle|^2 = |\langle T^{-1}x_m, x_m \rangle|^2 \le \|T^{-1}x_m\|^2 \|x_m\|^2$$

and

$$\|x_m\|^4 = \left|\langle x_m, x_m\rangle\right|^2 \le \sum_n \left|\langle x_m, x_n\rangle\right|^2 \le B\|x_m\|^2.$$

Consequently,

$$A \le \|x_m\|^2 \le B.$$

Hence, the sequence $\{x_n\}$ is bounded in norm, and $x \in H$ can be represented as

$$x = \sum_n \langle x, T^{-1}x_n\rangle x_n.$$

It remains to show that this representation is unique. If $x = \sum_n c_n x_n$, then

$$\langle x, T^{-1}x_m\rangle = \left\langle \sum_n c_n x_n, T^{-1}x_m \right\rangle = \sum_n c_n \langle x_n, T^{-1}x_m\rangle = c_m.$$

Thus, the sequence $\{x_n\}$ is a basis. Since the series converges unconditionally, the basis is unconditional.

The condition is sufficient.

We assume that $\{x_n\}$ is a bounded unconditional basis of H. Then, there exists an orthonormal basis $\{e_n\}$ and a topological isomorphism $T : H \to H$ such that $Te_n = x_n$ for all n. For $x \in H$, we have

$$\sum_n \left|\langle x, x_n\rangle\right|^2 = \sum_n \left|\langle x, Te_n\rangle\right|^2 = \sum_n \left|\langle T^* x, e_n\rangle\right|^2 = \|T^* x\|^2,$$

where T^* is the adjoint of T. But

$$\left\|T^{*-1}\right\|^{-1} \|x\| \le \|T^* x\| \le \|T^*\| \|x\|.$$

Hence, the sequence $\{x_n\}$ is a frame which is obviously an exact frame because it ceases to be a frame whenever any element is deleted from the sequence. This completes the proof.

4.6 Discrete Gabor Transforms and the Gabor Representation Problem

In many applications to physical and engineering problems, it is more important, at least from a computational viewpoint, to work with discrete transforms rather than continuous ones. In sampling theory, the sample points are defined by $\upsilon = m\omega_0$ and

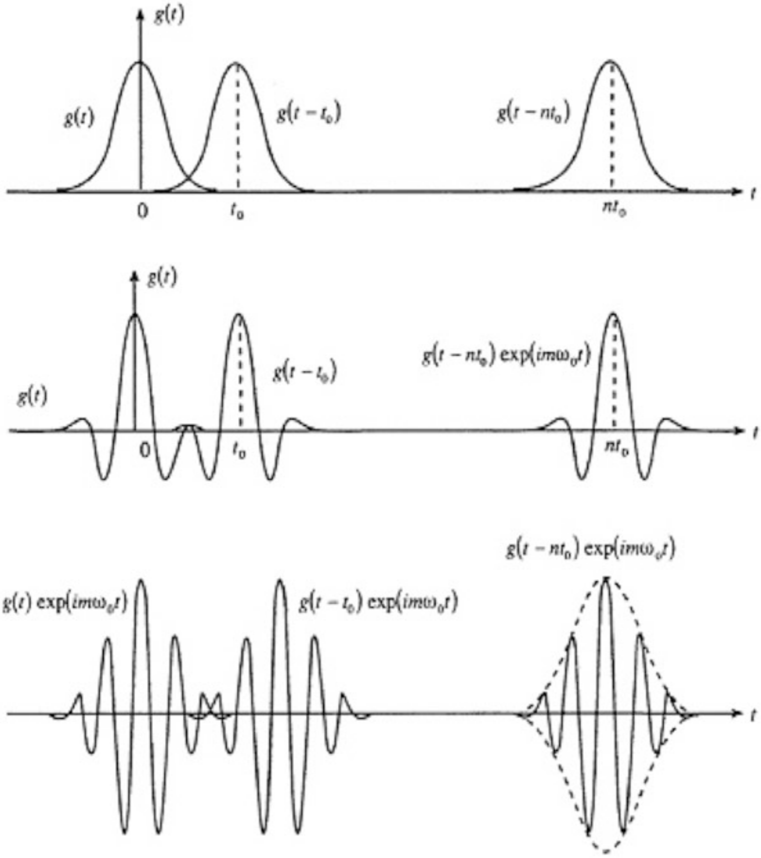

Fig. 4.6 The Gabor elementary functions $g_{m,n}(t)$

$\tau = nt_0$, where m, n are integers and t_0 and ω_0 are positive quantities. The *discrete Gabor functions* are defined by

$$g_{m,n}(t) = \exp(2\pi m\omega_0 t)\, g(t - nt_0) = M_{2\pi m\omega_0}\, T_{nt_0}\, g(t), \tag{4.6.1}$$

where $g \in L^2(\mathbb{R})$ is a fixed function and t_0 and ω_0 are the time shift and the frequency shift parameters, respectively. A typical set of Gabor functions is shown in Fig. 4.6.

These functions are also called the *Weyl–Heisenberg coherent states* which arise from translations and modulations of the Gabor window function (Fig. 4.6). From a physical point of view, these coherent states are of great interest and have several important applications in quantum mechanics. Following Gabor's analysis, various other functions have been introduced as window functions instead of the Gaussian function which was originally used by Gabor. In order to expand general functions (quantum mechanical states) with respect to states with minimum uncertainty, von Neumann (1945) introduced a set of coherent states on lattice constants $\omega_0 t_0 = h$ in the phase space with position and momentum as coordinates where h is the Planck constant. These states, associated with the *Weyl–Heisenberg group*, are in fact the same as used by Gabor. The time–frequency lattice with lattice constants $\omega_0 t_0 = 1$ is also called the *von Neumann lattice*.

Definition 4.6.1 (Discrete Gabor Transform). The discrete Gabor transform is defined by

$$\tilde{f}(m,n) = \int_{-\infty}^{\infty} f(t)\,\overline{g}_{m,n}(t)dt = \langle f, g_{m,n}\rangle. \tag{4.6.2}$$

The double series

$$\sum_{m,n=-\infty}^{\infty} \tilde{f}(m,n)\, g_{m,n}(t) = \sum_{m,n=-\infty}^{\infty} \langle f, g_{m,n}\rangle\, g_{m,n}(t) \tag{4.6.3}$$

is called the *Gabor series* of f.

It is of special interest to find the inverse of the discrete Gabor transform so that $f \in L^2(\mathbb{R})$ can be determined by the formula

$$\tilde{f}(mt_0, n\omega_0) = \int_{-\infty}^{\infty} f(t)\, g_{m,n}(t)\, dt = \langle f, \overline{g}_{m,n}\rangle. \tag{4.6.4}$$

The set of sample points $\{(mt_0, n\omega_0)\}_{m,n=-\infty}^{\infty}$ is called the *Gabor lattice*. The answer to the question of finding the inverse is in the affirmative if the set of functions $\{g_{m,n}(t)\}$ forms an orthonormal basis, or more generally, if the set is a frame for $L^2(\mathbb{R})$. A system $\{g_{m,n}(t)\} = \{M_{2\pi m\omega_0}\, T_{nt_0}\, g(t)\}$ is called a *Gabor frame* or *Weyl–Heisenberg frame* in $L^2(\mathbb{R})$ if there exist two constants $A, B > 0$ such that

$$A\|f\|^2 \le \sum_{m,n=-\infty}^{\infty} |\langle f, g_{m,n}\rangle|^2 \le B\|f\|^2 \tag{4.6.5}$$

holds for all $f \in L^2(\mathbb{R})$. For a Gabor frame $\{g_{m,n}(t)\}$, the *analysis operator* T_g is defined by

$$T_g f = \left\{\langle f, g_{m,n}\rangle\right\}_{m,n}, \tag{4.6.6}$$

and its *synthesis operator* T_g^* is defined by

$$T_g^* c_{m,n} = \sum_{m,n=-\infty}^{\infty} c_{m,n}\, g_{m,n}, \tag{4.6.7}$$

where $c_{m,n} \in \ell^2(\mathbb{Z})$. Both T_g and T_g^* are bounded linear operators and in fact are adjoint operators with respect to the inner product $\langle\, , \rangle$. The *Gabor frame operator* S_g is defined by $S_g = T_g^* T_g$. More explicitly,

$$S_g f = \sum_{m,n=-\infty}^{\infty} \langle f, g_{m,n} \rangle g_{m,n}. \tag{4.6.8}$$

If $\{g_{m,n}\}$ constitute a Gabor frame for $L^2(\mathbb{R})$, any function $f \in L^2(\mathbb{R})$ can be expressed as

$$f(t) = \sum_{m,n=-\infty}^{\infty} \langle f, g_{m,n} \rangle g_{m,n}^* = \sum_{m,n=-\infty}^{\infty} \langle f, g_{m,n}^* \rangle g_{m,n}, \tag{4.6.9}$$

where $\{g_{m,n}^*\}$ is called the *dual frame* given by $g_{m,n}^* = S_g^{-1} g_{m,n}$. Equation (4.6.9) provides an answer for constructing f from its Gabor transform $\langle f, g_{m,n} \rangle$ for a given window function g.

Finding the conditions on t_0, ω_0, and g under which the Gabor series of f determines f or converges to it is known as the *Gabor representation problem.* For an appropriate function g, the answer is positive provided that $0 < \omega_0 t_0 < 1$. If $0 < \omega_0 t_0 < 1$, the reconstruction is *stable* and g can have a good time and frequency localization. This is in contrast with the case when $\omega_0 t_0 = 1$, where the construction is *unstable* and g cannot have a good time and frequency localization. For the case when $\omega_0 t_0 > 1$, the reconstruction of f is, in general, impossible no matter how g is selected.

4.7 The Zak Transform and Time–Frequency Signal Analysis

Historically, the Zak transform (ZT), known as the *Weil-Brezin transform* in harmonic analysis, was introduced by Gelfand (1950) in his famous paper on eigenfunction expansions associated with Schrödinger operators with periodic potentials. This transform was also known as the *Gelfand mapping* in the Russian mathematical literature. However, Zak (1967, 1968) independently rediscovered it as the $k - q$ transform in solid state physics to study a quantum-mechanical representation of the motion of electrons in the presence of an electric or magnetic field. Although the Gelfand–Weil–Brezin–Zak transform seems to be a more appropriate name for this transform, there is a general consensus among scientists to name it as the Zak transform since Zak himself first recognized its deep significance and usefulness in a more general setting. In recent years, the Zak transform has been widely used in time–frequency signal analysis, in the coherent states representation in quantum field theory, and also in mathematical analysis of Gabor systems. In particular, the

Zak transform has also been useful for a study of the Gabor representation problem, where this transform has successfully been utilized to investigate the orthogonality and completeness of the Gabor frames in the critical case.

Definition 4.7.1 (The Zak Transform). The Zak transform, $\big(\mathscr{Z}_a f\big)(t,\omega)$, of a function $f \in L^2(\mathbb{R})$ is defined by the series

$$\big(\mathscr{Z}_a f\big)(t,\omega) = \sqrt{a} \sum_{n=-\infty}^{\infty} f(at+an)\exp(-2\pi i n\omega), \tag{4.7.1}$$

where $a(>0)$ is a fixed parameter and t and ω are real.

If f represents a signal, then its Zak transform can be treated as the joint time–frequency representation of the signal f. It can also be considered as the discrete Fourier transform of f in which an infinite set of samples in the form $f(at+an)$ is used for $n = 0, \pm 1, \pm 2, \ldots$. Without loss of generality, we set $a = 1$ so that we can write $\big(\mathscr{Z} f\big)(t,\omega)$ in the explicit form

$$\big(\mathscr{Z} f\big)(t,\omega) = F(t,\omega) = \sum_{n=-\infty}^{\infty} f(t+n)\exp(-2\pi i n\omega). \tag{4.7.2}$$

This transform satisfies the *periodic relation*

$$\big(\mathscr{Z} f\big)(t,\omega+1) = \big(\mathscr{Z} f\big)(t,\omega), \tag{4.7.3}$$

and the following *quasiperiodic relation*

$$\big(\mathscr{Z} f\big)(t+1,\omega) = \exp(2\pi i\omega)\big(\mathscr{Z} f\big)(t,\omega), \tag{4.7.4}$$

and therefore the Zak transform $\mathscr{Z} f$ is completely determined by its values on the unit square $S = [0,1] \times [0,1]$.

It is easy to prove that the Zak transform of f can be expressed in terms of the Zak transform of its Fourier transform $\hat{f}(\upsilon) = \mathscr{F}\{f(t)\}$ defined by (3.3.19b). More precisely,

$$\big(\mathscr{Z} f\big)(t,\omega) = \exp(2\pi i\omega t)\big(\mathscr{Z} \hat{f}\,\big)(\omega,-t). \tag{4.7.5}$$

To prove this result, we define a function g for fixed t and ω by

$$g(x) = \exp(-2\pi i\omega x)\, f(x+t).$$

Then, it follows that

$$\begin{aligned}\hat{g}(\upsilon) &= \int_{-\infty}^{\infty} g(x)\, \exp(-2\pi i x\upsilon)\, dx \\ &= \int_{-\infty}^{\infty} f(x+t)\, \exp\big\{-2\pi i x(\upsilon+\omega)\big\}\, dx\end{aligned}$$

$$= e^{2\pi i(\upsilon+\omega)t}\int_{-\infty}^{\infty} f(u)\,\exp\big\{-2\pi i(\upsilon+\omega)u\big\}\,du$$
$$= \exp\big\{2\pi i(\upsilon+\omega)t\big\}\hat{f}(\upsilon+\omega).$$

We next use the Poisson summation formula (3.7.7) in the form

$$\sum_{n=-\infty}^{\infty} g(n) = \sum_{n=-\infty}^{\infty} \hat{g}(2\pi n).$$

Or, equivalently,

$$\sum_{n=-\infty}^{\infty} f(t+n)\,\exp(-2\pi i\omega n) = \exp(2\pi i\omega t)\sum_{n=-\infty}^{\infty}\exp\big[2\pi i\,(2n\pi)t\big]\hat{f}(\omega+2\pi n)$$
$$= \exp(2\pi i\omega t)\sum_{m=-\infty}^{\infty}\hat{f}(\omega+m)\exp(2\pi i m t).$$

This gives the desired result (4.7.5).

The following results can be easily verified:

$$\big(\mathscr{Z}\mathscr{F}f\big)(\omega,t) = \exp(2\pi i\omega t)\big(\mathscr{Z}f\big)(-t,\omega), \tag{4.7.6}$$
$$\big(\mathscr{Z}\mathscr{F}^{-1}f\big)(\omega,t) = \exp(2\pi i\omega t)\big(\mathscr{Z}f\big)(-t,\omega). \tag{4.7.7}$$

If $g_{m,n}(t) = \exp(-2\pi i m t)g(t-n)$, then

$$\big(\mathscr{Z}g_{m,n}\big)(\omega,t) = \exp\big[-2\pi i(mt+n\omega)\big]\big(\mathscr{Z}g(\omega,t)\big). \tag{4.7.8}$$

We next observe that $L^2(S)$ is the set of all square integrable complex-valued functions F on the unit square S, that is,

$$\int_0^1\int_0^1 \big|F(t,\omega)\big|^2 dt\,d\omega < \infty.$$

It is easy to check that $L^2(S)$ is a Hilbert space with the inner product

$$\langle F,G\rangle = \int_0^1\int_0^1 F(t,\omega)\,\overline{G}(t,\omega)\,dt\,d\omega \tag{4.7.9}$$

and the norm

$$\|F\| = \left[\int_0^1\int_0^1 \big|F(t,\omega)\big|^2 dt\,d\omega\right]^{\frac{1}{2}}. \tag{4.7.10}$$

The set

$$\left\{M_{m,n} = M_{2\pi m, 2\pi n}(t,\omega) = \exp\left[2\pi i(mt + n\omega)\right]\right\}_{m,n=-\infty}^{\infty} \tag{4.7.11}$$

forms an orthonormal basis of $L^2(S)$.

Example 4.7.1. If

$$\phi_{m,n;a}(x) = \frac{1}{\sqrt{a}}\, T_{na}\, M_{2\pi m/a}\, \chi_{[0,a]}(x), \tag{4.7.12}$$

where $a > 0$, then

$$\left(\mathscr{Z}_a \phi_{m,n;a}\right)(t,\omega) = e_m(t)\, e_n(\omega), \tag{4.7.13}$$

where $e_k(t) = \exp(2\pi i k t)$.

We have

$$\begin{aligned}\phi_{m,n;a}(x) &= \frac{1}{\sqrt{a}} \exp\left[2\pi i m\left(\frac{x-na}{a}\right)\right] \chi_{[0,a]}(x-na)\\ &= \frac{1}{\sqrt{a}} \exp\left(\frac{2\pi i m x}{a}\right) \chi_{[na,(n+1)a]}(x).\end{aligned}$$

Thus, we obtain

$$\begin{aligned}\left(\mathscr{Z}_a \phi_{m,n;a}\right)(t,\omega) &= \sum_{k=-\infty}^{\infty} \exp\left[\frac{2\pi i m}{a}(at + ak)\right] \chi_{[na,na+a]}(at+ak)\\ &= \sum_{k=-\infty}^{\infty} e_m(t)\, e^{-2\pi i k\omega}\, \chi_{[n-k,n+1-k]}(t)\\ &= e_m(t)\, e_n(\omega).\end{aligned}$$

4.8 Basic Properties of Zak Transforms

1. (*Linearity*). The Zak transform is linear, that is, for any two constants a, b,

$$\left[\mathscr{Z}(af + bg)\right](t,\omega) = a\left(\mathscr{Z}f\right)(t,\omega) + b\left(\mathscr{Z}g\right)(t,\omega). \tag{4.8.1}$$

2. (*Translation*). For any real a and integer m,

$$\left[\mathscr{Z}(T_a f)\right](t,\omega) = \left(\mathscr{Z}f\right)(t-a,\omega), \tag{4.8.2}$$

$$\left[\mathscr{Z}(T_{-m} f)\right](t,\omega) = \exp(2\pi i m\omega)\left(\mathscr{Z}f\right)(t,\omega), \tag{4.8.3}$$

3. (*Modulation*).

$$\left[\mathscr{Z}(M_b f)\right](t,\omega) = e^{ibt}\left(\mathscr{Z} f\right)\left(t, \omega - \frac{b}{2\pi}\right), \tag{4.8.4}$$

$$\left[\mathscr{Z}(M_{2\pi b} f)\right](t,\omega) = \exp(2\pi i b t)\left(\mathscr{Z} f\right)(t, \omega - b). \tag{4.8.5}$$

4. (*Translation and Modulation*).

$$\mathscr{Z}\left[M_{2\pi m} T_n f\right](t,\omega) = \exp\left[2\pi i(mt - n\omega)\right]\left(\mathscr{Z} f\right)(t,\omega). \tag{4.8.6}$$

5. (*Conjugation*).

$$\left(\mathscr{Z} \bar{f}\right)(t,\omega) = \overline{\left(\mathscr{Z} f\right)}(t,-\omega). \tag{4.8.7}$$

6. (*Symmetry*).
   (a) If f is an even function, then

$$\left(\mathscr{Z} f\right)(t,\omega) = \left(\mathscr{Z} f\right)(-t,-\omega). \tag{4.8.8}$$

   (b) If f is an odd function, then

$$\left(\mathscr{Z} f\right)(t,\omega) = -\left(\mathscr{Z} f\right)(-t,-\omega). \tag{4.8.9}$$

If f is a real and even function, it follows from (4.8.7) that

$$\left(\mathscr{Z} f\right)(t,\omega) = \overline{\left(\mathscr{Z} f\right)}(t,-\omega) = \left(\mathscr{Z} f\right)(-t,-\omega). \tag{4.8.10}$$

7. (*Inversion*). For $t, \omega \in \mathbb{R}$,

$$f(t) = \int_0^1 \left(\mathscr{Z} f\right)(t,\omega)\, d\omega, \tag{4.8.11}$$

$$\hat{f}(\omega) = \int_0^1 \exp(-2\pi i \omega t)\left(\mathscr{Z} f\right)(t,\omega)\, dt, \tag{4.8.12}$$

$$f(x) = \int_0^1 \exp(-2\pi i x t)\left(\mathscr{Z} \hat{f}\right)(t,x)\, dt. \tag{4.8.13}$$

8. (*Dilation*).

$$\left(\mathscr{Z} D_{\frac{1}{a}} f\right)(t,\omega) = \left(\mathscr{Z}_a f\right)\left(at, \frac{\omega}{a}\right). \tag{4.8.14}$$

9. (*Product and Convolution of Zak Transforms*).

Results (4.7.3) and (4.7.4) show that the Zak transform is not periodic in the two variables t and ω. The product of two Zak transforms is periodic in t and ω.

Proof. We consider the product

$$F(t,\omega) = \big(\mathscr{Z} f\big)(t,\omega)\overline{\big(\mathscr{Z} g\big)}(t,\omega) \tag{4.8.15}$$

and find from (4.7.4) that

$$\overline{\big(\mathscr{Z} g\big)}(t,\omega) = \exp(-2\pi i\omega)\big(\mathscr{Z} g\big)(t,\omega).$$

Therefore, it follows that

$$\begin{aligned}
F(t+1,\omega) &= \big(\mathscr{Z} f\big)(t,\omega)\overline{\big(\mathscr{Z} g\big)}(t,\omega) = F(t,\omega),\\
F(t,\omega+1) &= \big(\mathscr{Z} f\big)(t,\omega)\overline{\big(\mathscr{Z} g\big)}(t,\omega) = F(t,\omega).
\end{aligned}$$

These show that F is periodic in t and ω. Consequently, it can be expanded in a Fourier series on a unit square

$$F(t,\omega) = \sum_{m,n=-\infty}^{\infty} c_{m,n} \exp(2\pi i m t)\, \exp(2\pi i n\omega), \tag{4.8.16}$$

where

$$c_{m,n} = \int_0^1 \int_0^1 F(t,\omega) \exp(-2\pi i m t)\, \exp(-2\pi i n\omega)\, dt\, d\omega.$$

If we assume that the series involved are uniformly convergent, we can interchange the summation and integration to obtain

$$\begin{aligned}
c_{m,n} &= \int_0^1 \int_0^1 \left[\sum_{r=-\infty}^{\infty} f(t+r) \exp(-2\pi i r\omega)\right]\left[\sum_{s=-\infty}^{\infty} \bar{g}(t+s) \exp(2\pi i s\omega)\right]\\
&\quad \times \exp\big\{-2\pi i(mt+n\omega)\big\}\, dt\, d\omega\\
&= \int_0^1 \left[\sum_{r=-\infty}^{\infty} f(t+r)\right]\left[\sum_{s=-\infty}^{\infty} \bar{g}(t+s)\right] \exp(-2\pi i m t)\, dt\\
&\quad \times \int_0^1 \exp\big\{2\pi i\omega(s-n-r)\big\}\, d\omega\\
&= \int_0^1 \left[\sum_{r=-\infty}^{\infty} f(t+r)\, \bar{g}(t+n+r)\right] \exp(-2\pi i m t)\, dt
\end{aligned}$$

$$= \sum_{r=-\infty}^{\infty} \int_{r}^{r+1} f(x)\,\bar{g}(x+n) \exp\big\{-2\pi i m(x-r)\big\}\,dx$$

$$= \int_{-\infty}^{\infty} f(x)\,\bar{g}(x+n) \exp\left(-2\pi i m x\right)\,dx$$

$$= \Big\langle f(x), e^{2\pi i m x} g(x+n)\Big\rangle$$

$$= \big\langle f, M_{2\pi m}\,T_{-n} g\big\rangle.$$

Consequently, (4.8.16) becomes

$$\big(\mathscr{Z} f\big)(t,\omega)\overline{\big(\mathscr{Z} g\big)}(t,\omega) = \sum_{m,n=-\infty}^{\infty} \big\langle f, M_{2\pi m}\,T_{-n} g\big\rangle \exp\big\{2\pi i\,(mt+n\omega)\big\}. \tag{4.8.17}$$

This completes the proof.

Theorem 4.8.1. *Suppose H is a function of two real variables t and s satisfying the condition*

$$H\big(t+1, s+1\big) = H\big(t,s\big), \quad s,t \in \mathbb{R}, \tag{4.8.18}$$

and

$$h(t) = \int_{-\infty}^{\infty} H(t,s)\,f(s)\,ds, \tag{4.8.19}$$

where the integral is absolutely and uniformly convergent.

Then,

$$\big(\mathscr{Z} f\big)(t,\omega) = \int_0^1 \big(\mathscr{Z} f\big)(s,\omega)\,\Phi(t,s,\omega)\,ds, \tag{4.8.20}$$

where Φ is given by

$$\Phi(t,s,\omega) = \sum_{n=-\infty}^{\infty} H(t+n,s)\,\exp(-2\pi i n\omega), \qquad 0 \le t,s,\omega \le l. \tag{4.8.21}$$

Proof. It follows from the definition of the Zak transform of $h(t)$ that

$$\big(\mathscr{Z} h\big)(t,\omega) = \sum_{k=-\infty}^{\infty} h(t+k)\,e^{-2\pi i k\omega} = \sum_{k=-\infty}^{\infty} e^{-2\pi i k\omega} \int_{-\infty}^{\infty} H(t+k,s)\,f(s)\,ds$$

$$= \sum_{k=-\infty}^{\infty} e^{-2\pi i k\omega} \sum_{m=-\infty}^{\infty} \int_{m}^{m+1} H(t+k,s)\,f(s)\,ds$$

$$= \sum_{k=-\infty}^{\infty} e^{-2\pi i k\omega} \sum_{m=-\infty}^{\infty} \int_0^1 H(t+k, s+m) f(s+m)\, ds$$

$$= \int_0^1 \left[\sum_{k,m=-\infty}^{\infty} H(t+k, s+m)\, f(s+m)\, \exp(-2\pi i k\omega) \right] ds,$$

which is, due to (4.8.17)

$$= \int_0^1 \left[\sum_{k,m=-\infty}^{\infty} H(t+k-m, s)\, f(s+m)\, \exp(-2\pi i k\omega) \right] ds$$

$$= \int_0^1 \left[\sum_{m,n=-\infty}^{\infty} H(t+n, s)\, f(s+m)\, \exp\{-2\pi i (m+n)\omega\} \right] ds$$

$$= \int_0^1 \big(\mathscr{Z} f\big)(s, \omega)\, \Phi(t, s, \omega) ds. \tag{4.8.22}$$

This completes the proof.

In particular, if $H(t, s) = H(t - s)$,

$$\Phi(t, s, \omega) = \sum_{n=-\infty}^{\infty} H(t - s + n)\, \exp(-2\pi i n\omega) = \big(\mathscr{Z} H\big)(t - s, \omega).$$

Consequently, Theorem 4.8.1 leads to the following convolution theorem.

Theorem 4.8.2 (Convolution Theorem). *If*

$$h(t) = \int_{-\infty}^{\infty} H(t-s) f(s)\, ds = (H * f)(t),$$

then (4.8.20) reduces to the form

$$\big(\mathscr{Z} h\big)(t, \omega) = \int_0^1 \big(\mathscr{Z} H\big)(t-s)\, \big(\mathscr{Z} f\big)(s, \omega)\, ds = \mathscr{Z}(H * f)(t, \omega). \tag{4.8.23}$$

Example 4.8.1. If $H(t) = \sum_{k=-\infty}^{\infty} a_k\, \delta(t-k)$, then

$$\mathscr{Z}(H * f)(t, \omega) = A(\omega)\big(\mathscr{Z} f\big)(t, \omega), \tag{4.8.24}$$

where

$$A(\omega) = \sum_{k=-\infty}^{\infty} a_k \exp(-2\pi i k\omega).$$

Clearly,

$$\begin{aligned}
\mathscr{Z}(H * f)(t,\omega) &= \mathscr{Z}\left[\int_{-\infty}^{\infty} H(t-s)\,f(s)\,ds\right](t,\omega)\\
&= \mathscr{Z}\left[\sum_{k=-\infty}^{\infty} a_k \int_{-\infty}^{\infty} \delta(t-s-k)\,f(s)\,ds\right](t,\omega)\\
&= \mathscr{Z}\left[\sum_{k=-\infty}^{\infty} a_k\, f(t-k)\right](t,\omega)\\
&= \sum_{k=-\infty}^{\infty} a_k \sum_{n=-\infty}^{\infty} f(t+n-k)\exp(-2\pi i n\omega)\\
&= \sum_{k=-\infty}^{\infty} a_k \sum_{m=-\infty}^{\infty} f(t+m)\exp\big\{-2\pi i\omega(m+k)\big\}\\
&= A(\omega)\big(\mathscr{Z} f\big)(t,\omega).
\end{aligned}$$

Theorem 4.8.3. *The Zak transformation is a unitary mapping from* $L^2(\mathbb{R})$ *to* $L^2(S)$.

Proof. It follows from the definition of the inner product (4.7.9) in $L^2(S)$ that

$$\begin{aligned}
\langle \mathscr{Z}_a f, \mathscr{Z}_a g\rangle &= a\int_0^1\int_0^1\left[\sum_{n=-\infty}^{\infty} f(at+an)\,e^{-2\pi i n\omega}\right]\left[\sum_{m=-\infty}^{\infty}\overline{g}(at+am)\,e^{2\pi i m\omega}\right]dt\,d\omega\\
&= a\int_0^1\left[\sum_{n=-\infty}^{\infty} f(at+an)\,\overline{g}(at+an)\right]dt\\
&= \sum_{n=-\infty}^{\infty}\int_{na}^{(n+1)a} f(x)\,\overline{g}(x)\,dx\\
&= \int_{-\infty}^{\infty} f(x)\,\overline{g}(x)\,dx = \langle f, g\rangle.
\end{aligned}\tag{4.8.25}$$

In particular, if $f = g$, we obtain from (4.8.25) that

$$\big\|\mathscr{Z}_a f\big\|^2 = \big\|f\big\|^2. \tag{4.8.26}$$

This means that the Zak transform is an isometry from $L^2(\mathbb{R})$ to $L^2(S)$.

Further, Example 4.7.1 shows that $\big\{\phi_{m,na}(x)\big\}_{m,n=-\infty}^{\infty}$ is an orthonormal basis of $L^2(\mathbb{R})$. Hence, the Zak transform is a one-to-one mapping of an orthonormal basis of $L^2(\mathbb{R})$ onto an orthonormal basis of $L^2(S)$. This proves the theorem.

4.9 Applications of Zak Transforms and the Balian–Low Theorem

It has already been mentioned that the Zak transform plays a major role in the study of the Gabor representation problem in signal analysis and the coherent states representation in quantum physics. Furthermore, the Zak transform is particularly useful in proving the Balian–Low theorem (BLT) which is also a fundamental result in time–frequency analysis. For a detailed investigation of these problems, we need the following results.

If $t_0, \omega_0 > 0,\ g \in L^2(\mathbb{R})$, and

$$g_{m,n}(t) = g_{m\omega_0,nt_0}(t) = M_{2\pi m\omega_0}\, T_{nt_0}\, g(t) = \exp(2\pi i m\omega_0 t)\, g(t - nt_0) \qquad (4.9.1)$$

is a Gabor system (or Weyl–Heisenberg system), then it is easy to verify that, if $\omega_0 t_0 = 1$,

$$\begin{aligned}\mathscr{Z}_{t_0}\big[g_{m,n}(t)\big](t,\omega) &= \exp\big\{2\pi i(mt - n\omega)\big\}\big(\mathscr{Z}_{t_0} g\big)(t,\omega)\\ &= e_m(t)\, e_{-n}(\omega)\big(\mathscr{Z}_{t_0} g\big)(t,\omega),\end{aligned} \qquad (4.9.2)$$

where $e_k(t) = \exp(2\pi i k t)$.

Furthermore, if $\big\{g_{m,n}(t)\big\}$ is a frame in $L^2(\mathbb{R})$, then the frame operator S is given by

$$Sf = \sum_{m,n=-\infty}^{\infty} \big\langle f, g_{m,n}\big\rangle g_{m,n}, \qquad (4.9.3)$$

where $f \in L^2(\mathbb{R})$.

Theorem 4.9.1. *If $t_0, \omega_0 > 0,\ g \in L^2(\mathbb{R})$, and $\big\{g_{m,n}\big\}_{m,n=-\infty}^{\infty}$ is a frame in $L^2(\mathbb{R})$, then its dual frame $\big\{S^{-1} g_{m,n}\big\}_{m,n=-\infty}^{\infty}$ is also generated by one single function, More precisely,*

$$S^{-1} g_{m,n} = g^*_{m,n}, \qquad (4.9.4)$$

where $g^* = S^{-1} g$.

Proof. For any $f \in L^2(\mathbb{R})$ and fixed integer k, we have

$$\begin{aligned}S\big(T_{kt_0} f\big)(t) &= \sum_{m,n=-\infty}^{\infty} \big\langle T_{kt_0} f, g_{m,n}\big\rangle g_{m,n}(t)\\ &= \sum_{m,n=-\infty}^{\infty} \exp(-2\pi i m\omega_0 k t_0)\big\langle f, g_{m,n-k}\big\rangle g_{m,n}(t)\end{aligned}$$

$$= \sum_{m,n=-\infty}^{\infty} \exp(-2\pi i m\omega_0 k t_0)\langle f, g_{m,n}\rangle g_{m,n+k}(t)$$

$$= \sum_{m,n=-\infty}^{\infty} \langle f, g_{m,n}\rangle \exp\{2\pi i m\omega_0(t-kt_0)\} g(t-nt_0-kt_0)$$

$$= \sum_{m,n=-\infty}^{\infty} \langle f, g_{m,n}\rangle T_{kt_0}\big[\exp(2\pi i m\omega_0 t)\, g(t-nt_0)\big]$$

$$= T_{kt_0}\big(Sf(t)\big), \tag{4.9.5}$$

in which $T_{kt_0}\exp(2\pi i m\omega_0 t) = \exp(2\pi i m\omega_0 t)$ is used. This shows that S commutes with T_{kt_0}.

Similarly, S commutes with modulation operator $M_{2\pi k\omega_0}$ and hence

$$S\big(M_{2\pi k\omega_0}T_{st_0}f\big) = M_{2\pi k\omega_0}T_{st_0}\big(Sf\big). \tag{4.9.6}$$

Consequently,

$$S^{-1}\big(M_{2\pi k\omega_0}T_{st_0}f\big) = M_{2\pi k\omega_0}T_{st_0}f^*, \tag{4.9.7}$$

where $f^* = S^{-1}f$. Putting $f = g$ in (4.9.7) gives

$$S^{-1}\big(g_{m,n}\big) = S^{-1}\big(M_{2\pi m\omega_0}T_{nt_0}g\big) = M_{2\pi m\omega_0}T_{nt_0}S^{-1}g = M_{2\pi m\omega_0}T_{nt_0}g^* = g^*_{m,n}.$$

This completes the proof.

Remark. The elements of the dual frame $\{g^*_{m,n}\}$ are generated by a single function g^*, analogously to $g_{m,n}$. To compute the dual system, it is necessary to find the *dual atom* $g^* = S^{-1}g$ and compute all other elements $g^*_{m,n}$ of the dual frame by modulation and translation.

Some important properties of the Gabor system $\{g_{m,n}\}$ for $\omega_0 t_0 = 1$ are given by the following:

Theorem 4.9.2. *If $t_0, \omega_0 > 0$ such that $\omega_0 t_0 = 1$ and $g \in L^2(\mathbb{R})$, then the following statements are equivalent:*

(i) There exist two constants A and B such that

$$0 < A \le \left|\big(\mathscr{Z}_{t_0}\, g\big)(t,\omega)\right|^2 \le B < \infty.$$

(ii) The Gabor system $\{g_{m,n}(t) = \exp(2\pi i\, m\omega_0 t)g(t-nt_0)\}_{m,n=-\infty}^{\infty}$ is a frame in $L^2(\mathbb{R})$ with the frame bounds A and B.

(iii) The system $\{g_{m,n}(t)\}_{m,n=-\infty}^{\infty}$ is an exact frame in $L^2(\mathbb{R})$ with the frame bounds A and B.

If any of the above statements are satisfied, then there exists a unique representation of any $f \in L^2(\mathbb{R})$ in the form

$$f(t) = \sum_{m,n=-\infty}^{\infty} a_{m,n}\, g_{m,n}(t) = \sum_{m,n=-\infty}^{\infty} \langle f, g_{m,n}^{*} \rangle g_{m,n}(t), \tag{4.9.8}$$

where

$$a_{m,n} = \langle f, g_{m,n}^{*} \rangle = \int_0^1 \int_0^1 \frac{\left(\mathscr{Z}_{t_0} f\right)(t,\omega)}{\left(\mathscr{Z}_{t_0} g\right)(t,\omega)}\, e_{-m}(t)\, e_n(\omega)\, dt\, d\omega. \tag{4.9.9}$$

Proof. We first show that (i) implies (ii). Since Theorem 4.8.3 asserts that the Zak transformation is a unitary mapping from $L^2(\mathbb{R})$ onto $L^2(S)$, it suffices to prove that $\left\{\left(\mathscr{Z}_{t_0}\, g_{m,n}\right)(t,\omega)\right\}_{m,n=-\infty}^{\infty}$ is a frame in $L^2(S)$. Let $h \in L^2(S)$. Since $\left(\mathscr{Z}_{t_0}\, g\right)$ is bounded, $h\,\overline{\left(\mathscr{Z}_{t_0}\, g\right)} \in L^2(S)$, and hence, it follows from (4.9.2) that

$$\langle h, \mathscr{Z}_{t_0}\, g_{m,n} \rangle = \langle h, e_m(t)\, e_{-n}(\omega) \mathscr{Z}_{t_0}\, g \rangle = \left\langle h\,\overline{\left(\mathscr{Z}_{t_0}\, g\right)}, e_m(t)\, e_{-n}(\omega) \right\rangle. \tag{4.9.10}$$

Since $\{e_{m\omega_0}\, e_{-nt_0}\}$ is an orthonormal basis of $L^2(S)$, the Parseval relation implies that

$$\sum_{m,n=-\infty}^{\infty} \left|\langle h, \mathscr{Z}_{t_0}\, g_{m,n} \rangle\right|^2 = \left\| h\,\overline{\left(\mathscr{Z}_{t_0}\, g\right)} \right\|^2. \tag{4.9.11}$$

Combining this equality with the inequalities

$$A\|h\|^2 \le \left\| h\,\overline{\left(\mathscr{Z}_{t_0}\, g\right)} \right\|^2 \le B\|h\|^2$$

leads to the result

$$A\|h\|^2 \le \sum_{m,n=-\infty}^{\infty} \left|\langle h, \mathscr{Z}_{t_0}\, g_{m,n} \rangle\right|^2 \le B\|h\|^2.$$

This shows that $\left(\mathscr{Z}_{t_0}\, g_{m,n}\right)(t,\omega)$ is a frame in $L^2(S)$.

We next show that (ii) implies (i). If (ii) holds, then $\left\{e_m(t)\, e_{-n}(\omega)\left(\mathscr{Z}_{t_0}\, g\right)\right\}$ is a frame in $L^2(S)$ with frame bounds A and B. Hence, for any $h \in L^2(S)$, we must have

$$A\|h\|^2 \le \sum_{m,n=-\infty}^{\infty} \left|\langle h, e_m(t)\, e_{-n}(\omega)\left(\mathscr{Z}_{t_0}\, g\right)\right|^2 \le B\|h\|^2. \tag{4.9.12}$$

It follows from (4.9.10) and (4.9.11) that

$$\sum_{m,n=-\infty}^{\infty} \left|\left\langle h, e_m(t)\, e_{-n}(\omega)\left(\mathscr{Z}_{t_0}\, g\right)\right\rangle\right|^2 = \sum_{m,n=-\infty}^{\infty} \left|\left\langle h\, \overline{\left(\mathscr{Z}_{t_0}\, g\right)}, e_m(t)\, e_{-n}(\omega)\right\rangle\right|^2$$

$$= \left\| h\, \overline{\left(\mathscr{Z}_{t_0}\, g\right)}\right\|^2. \tag{4.9.13}$$

Combining (4.9.12) and (4.9.13) together gives

$$A\|h\|^2 \le \left\| h\, \overline{\left(\mathscr{Z}_{t_0}\, g\right)}\right\|^2 \le B\|h\|^2$$

which implies (i).

Next, we prove that (ii) implies (iii). Suppose (ii) is satisfied. Then $\left\{e_{m\omega_0}(t)\, e_{-nt_0}(\omega)\left(\mathscr{Z}_{t_0}\, g\right)\right\}$ represents a frame in $L^2(S)$. But (i) implies $\left(\mathscr{Z}_{t_0}\, g\right)$ is bounded. Hence the mapping $F : L^2(S) \to L^2(S)$ defined by

$$F(h) = F\left(\mathscr{Z}_{t_0}\, g\right), \qquad h \in L^2(S) \tag{4.9.14}$$

is a topological isomorphism that maps the orthonormal basis $\{e_m\, e_{-n}\}$ onto $\left\{\left(\mathscr{Z}_{t_0}\, g_{m,n}\right)(t,\omega)\right\}$. Thus, $\left\{\left(\mathscr{Z}_{t_0}\, g_{m,n}\right)(t,\omega)\right\}$ is a Riesz basis on $L^2(S)$ and hence so is $\left\{g_{m,n}(t,\omega)\right\}$ in $L^2(\mathbb{R})$. In view of the fact that $\left\{g_{m,n}(t,\omega)\right\}$ is a Riesz basis in $L^2(\mathbb{R})$, $\left\{g_{m,n}(t,\omega)\right\}$ is an exact frame for $L^2(\mathbb{R})$.

Finally, that (iii) implies (ii) is obvious. To prove (4.9.9), we first prove that

$$\mathscr{Z}_{t_0}(Sf) = \left(\mathscr{Z}_{t_0}\, f\right)\left|\left(\mathscr{Z}_{t_0}\, g\right)\right|^2, \tag{4.9.15}$$

where S is the frame operator associated with the frame $\left\{g_{m,n}(x)\right\}$. Since $\{e_m(t)\, e_{-n}(\omega)\}$ is an orthonormal basis for $L^2(S)$, it follows from (4.8.24) and (4.9.2) that

$$\begin{aligned}
\mathscr{Z}_{t_0}(Sf) &= \mathscr{Z}_{t_0}\left(\sum_{m,n=-\infty}^{\infty} \langle f, g_{m,n}\rangle g_{m,n}\right) \\
&= \left(\mathscr{Z}_{t_0} g\right) \sum_{m,n=-\infty}^{\infty} \left\langle f, g_{m,n}\right\rangle e_m(t)\, e_{-n}(\omega) \\
&= \left(\mathscr{Z}_{t_0} g\right) \sum_{m,n=-\infty}^{\infty} \left\langle \mathscr{Z}_{t_0} f, \mathscr{Z}_{t_0} g_{m,n}\right\rangle e_m(t)\, e_{-n}(\omega), \quad \text{by (4.8.25)} \\
&= \left(\mathscr{Z}_{t_0} g\right) \sum_{m,n=-\infty}^{\infty} \langle \mathscr{Z}_{t_0} f, \mathscr{Z}_{t_0} g\, e_m(t)\, e_{-n}(\omega)\rangle\; e_m(t)\, e_{-n}(\omega)
\end{aligned}$$

$$= \left(\mathscr{Z}_{t_0} g\right) \sum_{m,n=-\infty}^{\infty} \left\langle \mathscr{Z}_{t_0} f \, \overline{\mathscr{Z}_{t_0} g}, e_m(t)\, e_{-n}(\omega) \right\rangle e_m(t)\, e_{-n}(\omega)$$

$$= \left(\mathscr{Z}_{t_0} f\right) \left|\left(\mathscr{Z}_{t_0} g\right)\right|^2 .$$

This proves the result (4.9.15).

If we replace f by $S^{-1} f$ in (4.9.15), we obtain

$$\mathscr{Z}_{t_0}\left(S^{-1} f\right) = \frac{\left(\mathscr{Z}_{t_0} f\right)}{\left|\left(\mathscr{Z}_{t_0} g\right)\right|^2}, \tag{4.9.16}$$

which is, by putting $f = g$,

$$\mathscr{Z}_{t_0} g^* = \frac{1}{\overline{\left(\mathscr{Z}_{t_0} g\right)}}, \qquad g^* = S^{-1} g. \tag{4.9.17}$$

In view of (4.8.25), (4.9.2), (4.9.17), and Theorem 4.9.1, it turns out that

$$\begin{aligned} a_{m,n} &= \left\langle f, S^{-1} g_{m,n} \right\rangle = \left\langle f, g^*_{m,n} \right\rangle = \left\langle \mathscr{Z}_{t_0} f, \mathscr{Z}_{t_0} g^*_{m,n} \right\rangle \\ &= \left\langle \mathscr{Z}_{t_0} f, e_m(t)\, e_{-n}(\omega) \mathscr{Z}_{t_0} g^* \right\rangle \\ &= \left\langle \mathscr{Z}_{t_0} f, \frac{e_m(t)\, e_{-n}(\omega)}{\left(\mathscr{Z}_{t_0} g\right)} \right\rangle = \left\langle \frac{\mathscr{Z}_{t_0} f}{\mathscr{Z}_{t_0} g}, e_m(t)\, e_{-n}(\omega) \right\rangle \end{aligned}$$

which gives (4.9.9).

The *Gabor representation problem* can be stated as follows. Given $g \in L^2(\mathbb{R})$ and two real numbers t_0 and ω_0 different from zero, is it possible to represent any $f \in L^2(\mathbb{R})$ in the series form

$$f(t) = \sum_{m,n=-\infty}^{\infty} a_{m,n}\, g_{m,n}(t), \tag{4.9.18}$$

where $g_{m,n}$ is the Gabor system defined by (4.9.1) and $a_{m,n}$ are constants? Under what conditions is this representation unique?

Evidently, the above representation is possible, if the Gabor system $\{g_{m,n}\}$ forms an orthonormal basis or a frame in $L^2(\mathbb{R})$, and the uniqueness of the representation depends on whether the Gabor functions form a complete set in $L^2(\mathbb{R})$. The Zak transform is used to study this representation problem with two positive real numbers t_0 and ω_0 with $\omega_0 t_0 = 1$. We also use the result (4.9.2).

Theorem 4.9.3. *If t_0 and ω_0 are two positive real numbers with $\omega_0 t_0 = 1$ and $g \in L^2(\mathbb{R})$, then*

(i) the Gabor system $\{g_{m,n}\}$ is an orthonormal basis of $L^2(\mathbb{R})$ if and only if $\left|\left(\mathscr{Z}_{t_0} g\right)\right| = 1$ almost everywhere.
(ii) the Gabor system $\{g_{m,n}\}$ is complete in $L^2(\mathbb{R})$ if and only if $\left|\left(\mathscr{Z}_{t_0} g\right)\right| > 0$ almost everywhere.

Proof. (i) It follows from (4.8.25), (4.9.2), and Theorem 4.8.3 that

$$\langle g_{k,\ell}, g_{m,n}\rangle = \langle \mathscr{Z}_{t_0} g_{k,\ell}, \mathscr{Z}_{t_0} g_{m,n}\rangle = \int_0^1 \int_0^1 e_k(t)\, e_{-\ell}(\omega)\, \bar{e}_m(t)\, \bar{e}_{-n}(\omega) \left|\left(\mathscr{Z}_{t_0} g\right)\right|^2 dt\, d\omega.$$

This shows that the set $\{\mathscr{Z}_{t_0} g_{m,n}\}$ is an orthonormal basis in $L^2(\mathbb{R})$ if and only if $\left|\left(\mathscr{Z}_{t_0} g\right)\right| = 1$ almost everywhere.

An argument similar to above gives

$$\begin{aligned} \langle f, g_{m,n}\rangle &= \langle \mathscr{Z}_{t_0} f, \mathscr{Z}_{t_0} g_{m,n}\rangle = \langle \mathscr{Z}_{t_0} f, e_m(t)\, e_{-n}(\omega) \mathscr{Z}_{t_0} g\rangle \\ &= \langle \mathscr{Z}_{t_0} f\, \overline{\mathscr{Z}_{t_0} g}, e_m(t)\, e_{-n}(\omega)\rangle. \end{aligned} \tag{4.9.19}$$

This implies that $\{g_{m,n}\}$ is complete in $L^2(\mathbb{R})$ if and only if $\mathscr{Z}_{t_0} g \neq 0$ almost everywhere.

The answer to the Gabor representation problem can be summarized as follows.

The properties of the Gabor system $\{g_{m,n}\}$ are related to the density of the rectangular lattice $\Lambda = \{n t_0, m \omega_0\} = n\mathbb{Z} \times m\mathbb{Z}$ in the time–frequency plane. Small values of t_0, ω_0 correspond to a high density for Λ, whereas large values of t_0, ω_0 correspond to low density. Thus, it is natural to classify Gabor systems according to the following sampling density of the time–frequency lattice.

Case (i) (Oversampling). A Gabor system $\{g_{m,n}\}$ can be a frame where $0 < \omega_0 t_0 < 1$. In this case, frames exist with excellent time–frequency localization.

Case (ii) (Critical Sampling). This critical case corresponds to $\omega_0 t_0 = 1$, and there is a frame, and orthonormal basis exist, but g has bad localization properties either in time or in the frequency domain. More precisely, this case leads to the celebrated result in the time–frequency analysis which is known as the *BLT*, originally and independently stated by Balian (1981) and Low (1985) as follows.

Theorem 4.9.4 (Balian–Low). *If a Gabor system $\{g_{m,n}\}$ defined by (4.6.1) with $\omega_0 t_0 = 1$ forms an orthonormal basis in $L^2(\mathbb{R})$, then either $\int_{-\infty}^{\infty} \left|t g(t)\right|^2 dt$ or $\int_{-\infty}^{\infty} \left|\omega \hat{g}(\omega)\right|^2 d\omega$ must diverge, or equivalently,*

$$\int_{-\infty}^{\infty} \left|t g(t)\right|^2 dt \int_{-\infty}^{\infty} \left|\omega \hat{g}(\omega)\right|^2 d\omega = \infty. \tag{4.9.20}$$

The condition $\omega_0 t_0 = 1$ *associated with the density* $\Lambda = 1$ *can be interpreted as a Nyquist phenomenon for the Gabor system. In this critical situation, the time–frequency shift operators that are used to build a coherent frame commute with each other.*

For an elegant proof of the BLT using the Zak transform, we refer the reader to Daubechies (1992) or Benedetto and Frazier (1994).

Case (iii) (Undersampling). In this case, $\omega_0 t_0 > 1$. There is no frame of the form $\{g_{m,n}\}$ for any choice of the Gabor window function g. In fact, $\{g_{m,n}\}$ is incomplete in the sense that there exist $f \in L^2(\mathbb{R})$ such that $\langle f, g_{m,n}\rangle = 0$ for all m, n but $f \neq 0$.

These three cases can be represented by three distinct regions in the $t_0 - \omega_0$ plane, where the critical curve $\omega_0 t_0 = 1$ represents a *hyperbola* which separates the region $\omega_0 t_0 < 1$, where an exact frame exists with an excellent time–frequency localization from the region $\omega_0 t_0 > 1$ with no frames.

There exist many examples for g so that $\{g_{m,n}\}$ is a frame or even an orthonormal basis for $L^2(\mathbb{R})$. We give two examples of functions for which the family $\{M_{m\omega_0} T_{nt_0} g\}$ represents an orthonormal basis.

Example 4.9.1 (Characteristic Function). This function $g(t) = \chi_{[0,1]}(t)$ is defined by

$$g(t) = \begin{cases} 1, & 0 \leq t \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

Clearly,

$$\int_{-\infty}^{\infty} \left|\omega\, \hat{g}(\omega)\right|^2 d\omega = \infty.$$

Example 4.9.2 (Sine Function). In this case,

$$g(t) = \sin c(t) = \frac{\sin \pi t}{\pi t}.$$

Evidently, $\displaystyle\int_{-\infty}^{\infty} \left|t\, g(t)\right|^2 dt = \infty.$

Thus, these examples lead to systems with bad localization properties in either time or frequency. Even if the orthogonality requirement is dropped, we cannot construct Riesz bases with good time–frequency localization properties for the critical case $\omega_0 t_0 = 1$. This constitutes the contents of the BLT which describes one of the fundamental features of Gabor wavelet analysis.

4.10 Exercises

1. If $g(x) = \dfrac{1}{\sqrt{4\pi a}} \exp\left(-\dfrac{x^2}{4a}\right)$ is a Gaussian window, show that

(a) $\displaystyle\int_{-\infty}^{\infty} \tilde{f}_g(t, \omega)\,dt = \hat{f}(\omega), \quad \omega \in \mathbb{R}$

Give a significance of result l(a).

(b) $\hat{g}(\upsilon) = \exp(-a\upsilon^2)$.

2. Suppose $g_{t,\omega}(\tau) = g(\tau - t)\exp(i\omega\tau)$ where g is a Gaussian window defined in Exercise 1, show that

(a) $\hat{g}_{t,\omega}(\upsilon) = \exp\big[-i(\upsilon - \omega)t - a(\upsilon - \omega)^2\big]$.

(b) $\tilde{f}_g(t, \omega) = \dfrac{1}{2\pi}\langle \hat{f}, \hat{g}_{t,\omega}\rangle = \dfrac{1}{2\pi}\, e^{i\omega t}\, \bar{\hat{\tilde{f}}}_{\hat{g}}(t, \omega)$.

3. For the Gaussian window defined in Exercise 1, introduce

$$\sigma_t^2 = \frac{1}{\|g\|_2}\left\{\int_{-\infty}^{\infty} \tau^2 g^2(\tau)\, d\tau\right\}^{\frac{1}{2}}.$$

Show that the radius of the window function is $\sqrt{a}$ and the width of the window is twice the radius.

4. If $e_1 = (1, 0), e_2 = \left(-\dfrac{1}{2}, \dfrac{\sqrt{3}}{2}\right), e_3 = \left(-\dfrac{1}{2}, -\dfrac{\sqrt{3}}{2}\right)$ represent a set of vectors, show that, for any vector $x = (x_1, x_2)$,

$$\sum_{n=1}^{3} \left|\langle x, e_n\rangle\right|^2 = \frac{3}{2}\,\|x\|^2.$$

Hence, show that $\{e_i\}$ is a tight frame and $e_n^* = \dfrac{2}{3}\, e_n$.

5. If $e_1 = (1, 0), e_2 = (0, 1), e_3 = (-1, 0), e_4 = (0, -1)$ form a set of vectors, show that, for any vector x$x = (x_1, x_2)$,

$$\sum_{n=1}^{4} \left|\langle x, e_n\rangle\right|^2 = 2\,\|x\|^2$$

and

$$x = \sum_{k=1}^{4} \frac{1}{2}\langle x, x_k\rangle x_k.$$

6. If $e_1 = (1, 0), e_2 = \left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$ and $e_3 = \left(-\frac{1}{2}, -\frac{3}{2}\right)$ represent a set of vectors and $x = [x_1, x_2]^T$, show that

$$\sum_{n=1}^{3} \left|\langle x, e_n\rangle\right|^2 = \frac{1}{2}\left(x_1^2 + 5x_2^2\right),$$

and

$$\frac{1}{2}\left(x_1^2 + x_2^2\right) \le \sum_{n=1}^{3} \left|\langle x, e_n\rangle\right|^2 \le \frac{5}{2}\left(x_1^2 + x_2^2\right).$$

7. Show that the set of elements $\{e_n\}$ in a Hilbert space $\mathbb{C}^2$ forms a tight frame.

8. If g is a continuous function on $\mathbb{R}$ and if there exists an $\varepsilon > 0$ such that $|g(x)| \le A\big(1 + |x|\big)^{-1-\varepsilon}$, show that

$$g_{m,n}(x) = \exp(2\pi i m x) g(x - n)$$

cannot be a frame for $L^2(\mathbb{R})$.

9. Show that the marginals of the Zak transform are given by

$$\int_0^1 \big(\mathscr{Z} f\big)(t, \omega)\, d\omega = f(t),$$

$$\int_0^1 \exp(-2\pi i \omega t)\big(\mathscr{Z} f\big)(t, \omega)\, dt = \hat{f}(\omega).$$

10. If $f(t)$ is time-limited to $-a \le t \le a$ and band-limited to $-b \le \omega \le b$, where $0 \le a, b \le \frac{1}{2}$, then the following results hold:

$$\big(\mathscr{Z} f\big)(t, \omega) = f(\tau), \quad |\tau| \le \frac{1}{2}, \quad \omega \in \mathbb{R},$$

$$\big(\mathscr{Z} f\big)(t, \omega) = \exp(2\pi i \omega \tau)\hat{f}(\omega), \quad |\omega| \le \frac{1}{2}, \quad \tau \in \mathbb{R}.$$

Show that the second of the above results gives the Shannon's sampling formula

$$f(t) = \sum_{n=-\infty}^{\infty} \frac{\sin 2\pi b(n - t)}{\pi(n - t)}, \quad t \in \mathbb{R}.$$

11. If $g = \chi_{[0,1]}$, $g_{m,n}(x) = \exp(2\pi imx)g(x-n)$, where $m,n \in \mathbb{Z}$ is an orthonormal basis of $L^2(\mathbb{R})$, show that the first integral

$$\int_{-\infty}^{\infty} t\left|g(t)\right|^2 dt$$

in the BLT is finite, whereas the second integral

$$\int_{-\infty}^{\infty} \omega\left|\hat{g}(\omega)\right|^2 d\omega = \infty.$$

12. If $g(x) = \sin c(x) = \dfrac{\sin \pi x}{\pi x}$, $g_{m,n}(x) = \exp(2\pi imx)\, g(x-n)$ is an orthonormal basis of $L^2(\mathbb{R})$, show that the first integral in the BLT

$$\int_{-\infty}^{\infty} t\left|g(t)\right|^2 dt = \infty,$$

and the second integral in the BLT is finite.