Coherent states and wavelets, a contemporary panorama

Jean-Pierre Antoine

Abstract. In a first part, we review the general theory of coherent states (CS). Starting from the canonical CS introduced by Schrödinger in 1926 and rediscovered by Glauber, Klauder, Sudarshan in the 1960s, we proceed to the derivation of general CS from square integrable group representations and some of their generalizations: Perelomov CS, general square integrable covariant CS, nonlinear CS, Gazeau-Klauder CS, with a hint to their application in quantization.

Next, we turn to signal processing and note that two of the most familiar tools, namely, the Gabor transform and the wavelet transform, are special cases of CS, associated to the Weyl-Heisenberg group (which yields the canonical CS) and the affine group of the line, respectively. Then we review the properties of the wavelet transform, both in its continuous and its discrete versions, in one or two dimensions, emphasizing mostly the mathematical properties. We also consider its extension to higher dimensions, to more general manifolds (sphere, hyperboloid,...) and to the space-time context, for the analysis of moving objects.

Mathematics Subject Classification (2000). 22D10,42C40,62H35,68U10,81R30.

Keywords. Coherent states, wavelets, square integrable group representations, higher dimensional wavelets, wavelets on manifolds, signal processing, image processing.

1. Motivation

Coherent states (CS) were introduced in 1926 by Schrödinger for studying the quantum-to-classical transition, but quickly forgotten. They were rediscovered in the 1960s by Glauber, Klauder and Sudarshan in the context of quantum optics (coherent light from lasers) and it became readily clear that these so-called canonical CS, associated to a one-dimensional quantum harmonic oscillator, were in fact related to the Weyl-Heisenberg group. In 1972, Gilmore and Perelomov, independently, constructed CS associated to other groups, thus turning them into a problem in group representation theory. Since then, CS were considerably generalized, both mathematically and for applications into almost all domains of physics.

As for wavelets, they were invented around 1985 by Jean Morlet in the context of oil prospecting. Then it was soon realized, in collaboration with Alex Grossmann [21], that these continuous 1-D wavelets are in fact a particular case of CS, associated with the affine group of the line ('$ax + b$' group). There followed generalizations to higher dimensions and to various manifolds (two-sphere, two-sheeted hyperboloid, two-torus,...), then to more sophisticated mathematical objects, such as ridgelets, curvelets, shearlets, etc. In parallel, the theory exploded in a different direction, namely discrete wavelets, derived from multiresolution analysis. Nowadays, both the continuous and the discrete wavelet transforms have led to an immense number of applications, and they have become standard tools in signal/image processing. The aim of this chapter is to give an up-to-date panorama of the whole field, emphasizing its mathematical coherence. We mainly follow our textbooks [3, 10] and that of Gazeau [19], where plenty of applications are described in detail. In addition, all 'historical' papers on wavelets and their precursors have been collected in the compendium [22].

2. The pioneers : Canonical coherent states

2.1. Definitions

As we said above, the canonical CS are associated to a 1-D quantum harmonic oscillator. Starting from the canonical position and momentum operators Q, P, we introduce the operators

$$a = \frac{1}{\sqrt{2}}(Q + iP)\,, \quad a^\dagger = \frac{1}{\sqrt{2}}(Q - iP)\,, \quad [a, a^\dagger] = I. \tag{2.1}$$

Then the Hamiltonian of the system reads as $H = \frac{1}{2}(P^2 + Q^2) = a^\dagger a + \frac{1}{2} = N + \frac{1}{2}$ (we put $\hbar = 1$), where $N = a^\dagger a$ is the number operator, acting in the abstract Hilbert space $\mathcal{H}$. The spectrum of H is $\sigma(H) = \{n + \frac{1}{2},\ n = 0, 1, 2...\}$, the eigenstates $\{|n\rangle\}$ are nondegenerate and normalized, $\langle n|m\rangle = \delta_{mn}$, thus an orthonormal basis in $\mathcal{H}$.[1] In particular, the ground state satisfies the relation $a|\,0\rangle = 0$ and $a, a^\dagger$ are ladder operators :

$$a|n\rangle = \sqrt{n}\,|\,n-1\rangle, \qquad a^\dagger|n\rangle = \sqrt{n+1}\,|\,n+1\rangle.$$

Thus a is called an *annihilation* operator and $a^\dagger$ a *creation* operator. Indeed one can recover all states by acting successively on the ground state with the creation operator $a^\dagger$:

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}\,|\,0\,\rangle, \quad n = 1, 2, 3...$$

[1] We use here, and throughout the paper, the Dirac bra-ket notation familiar in quantum physics. In particular, our inner product $\langle\cdot|\cdot\rangle$ is linear in the *second* factor, so that $|\psi\rangle\langle\phi| := \psi \otimes \overline{\phi}$.

The problem is that these energy eigenstates are inadequate for describing the classical limit of quantum mechanics, $n \gg 1$, because they don't have a definite phase. Thus one needs other states and this is the reason why, in 1926, Schrödinger introduced the following ones, nowadays called *canonical CS*:

$$|\, z\,\rangle = e^{-\frac{1}{2}|z|^2} \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}}\,|\, n\rangle, \quad z \in \mathbb{C}. \tag{2.2}$$

Unfortunately, they were quickly forgotten, but were rediscovered in 1963 by Glauber, Klauder and Sudarshan in the context of quantum optics, for describing a coherent light beam, namely, a laser beam.

2.2. Basic properties

The canonical CS have four basic properties, that we now list.

(P1) The states $|z\rangle$ saturate the Heisenberg inequality:

$$\langle \Delta Q\rangle_z \; \langle \Delta P\rangle_z = \frac{1}{2}\hbar, \tag{2.3}$$

where, for $A = Q, P$, $\langle \Delta A\rangle_z := [\langle z|A^2|z\rangle - \langle z|A|z\rangle^2]^{1/2}$ is the variance (uncertainty) of the observable A in the state $|z\rangle$.

(P2) The states $|z\rangle$ are eigenvectors of the annihilation operator, with eigenvalue $z \in \mathbb{C}$ (thus they are uncountable):

$$a|z\rangle = z|z\rangle, \quad z \in \mathbb{C}. \tag{2.4}$$

(P3) The states $|z\rangle$ are obtained from the ground state $|0\rangle$ by a displacement operator

$$|z\rangle = D(z)|0\rangle, \text{ where } D(z) := e^{za^\dagger - \overline{z}a},\; z \in \mathbb{C}.$$

The operator $D(z)$ satisfies the relation $D(z)\,D(z') = e^{i\mathrm{Im}(z\overline{z'})}\,D(z+z')$.

(P4) The family $\{|z\rangle, z \in \mathbb{C}\}$ constitute an *overcomplete* (i.e., total) family of vectors in the Hilbert space $\mathcal{H}$. This property is encoded in the following *resolution of the identity*:

$$\int_{\mathbb{C}} |z\rangle\langle z| \, \frac{\mathrm{d}^2 z}{\pi} = I = \sum_{n=0}^{\infty} |n\rangle\langle n|, \tag{2.5}$$

where $\mathrm{d}^2 z := \mathrm{d}\,\mathrm{Re}\, z\; \mathrm{d}\,\mathrm{Im}\, z$ is the Lebesgue measure on the plane $\mathbb{C}$ and the integral is understood in the weak sense.

The properties (P1)–(P4) are equivalent for the canonical CS, in the sense that every one may be used as definition, but they lead to different generalizations.

Before going into these, let us discuss in more detail some consequences of the properties (P3) and (P4).

2.2.1. (P3) Connection with the Weyl-Heisenberg group. The three operators $\{a, a^\dagger, I\}$, with the commutation relation (2.1), generate a nilpotent Lie algebra. The corresponding Lie group, the *Weyl-Heisenberg* group $G_{\rm WH}$, plays a key role in nonrelativistic quantum mechanics. It consists of elements

$$\xi \equiv (\theta, z) := e^{\theta I + z a^\dagger - \overline{z} a} = e^{\theta I} D(z),\ \theta \in [0, 2\pi),\ z = 2^{-1/2}(q + ip) \in \mathbb{C}.$$

The center of $G_{\rm WH}$ is $Z = \{(\theta, 0)\} \simeq \mathbb{S}^1$. The group $G_{\rm WH}$ is unimodular, with (left and right) invariant measure

$$\mathrm{d}\mu(\xi) \,=\, (2\pi)^{-1}\, \mathrm{d}\theta\ \mathrm{d}q\ \mathrm{d}p \,=\, (2\pi)^{-1}\ \mathrm{d}\theta\ \mathrm{d}^2 z.$$

Given a unitary irreducible representation (UIR) U of $G_{\rm WH}$, its restriction to Z is $U(\theta, 0) = \exp(i\lambda\theta), \lambda \in \mathbb{Z}$. By von Neumann's uniqueness theorem, all UIRs of $G_{\rm WH}$ are unitarily equivalent and of the form $U^\lambda(\theta, z) = e^{i\lambda\theta}\, D^\lambda(z)$, where $D^\lambda(z) = D(z) = e^{z a^\dagger - \overline{z} a}$, acting in a Hilbert space $\mathcal{H}^\lambda$.

In addition, every representation U^λ ($\lambda \neq 0$) is *square integrable*:

$$\int_{G_{\rm WH}} |\langle U^\lambda(\xi)\phi|\phi\rangle|^2\ \mathrm{d}\mu(\xi) = \int_{G_{\rm WH}/Z} |\langle D^\lambda(z)\phi|\phi\rangle|^2\, \mathrm{d}^2 z \,<\, \infty,\ \forall\, \phi \in \mathcal{H}^\lambda.$$

For U^λ and $|0\rangle \in \mathcal{H}^\lambda$, the orbit $\{|z\rangle = D^\lambda(z)|0\rangle, z \in \mathbb{C} \simeq G_{\rm WH}/Z\}$ is a *family of canonical CS*.

More generally, given any nonzero vector $\eta \in \mathcal{H}^\lambda$, the orbit $\{\eta_z = D^\lambda(z)\eta,\ z \in G_{\rm WH}/Z\}$ is a family of CS such that

$$\int_{G_{\rm WH}/Z} |\eta_z\rangle\langle\eta_z|\, \frac{\mathrm{d}^2 z}{\pi} := \int_{G_{\rm WH}/Z} \eta_z \otimes \overline{\eta}_z\, \frac{\mathrm{d}^2 z}{\pi} = I \quad \text{(weakly)}.$$

In that sense, one realizes that the construction of canonical CS is a problem in the theory of group representations!

In particular, in the Schrödinger realization of quantum mechanics, one has $\mathcal{H} = L^2(\mathbb{R}, \mathrm{d}x)$ and

$$\begin{aligned} D(q,p)f)(x) &= e^{iqp/2}\, e^{ipx}\, f(x - q),\ x \in \mathbb{R}, \\ U^\lambda(\theta, q, p) &= e^{i\lambda\theta} D(q, p), \end{aligned}$$

thus U^λ is a UIR of $G_{\rm WH}$ in $L^2(\mathbb{R}, \mathrm{d}x)$.

2.2.2. (P4) Functional analysis of canonical CS. Condition (P4) implies that

$$|\psi\rangle = \int_{\mathbb{C}} |\overline{z}\rangle\langle\overline{z}|\psi\rangle\, \mathrm{d}\mu(\overline{z}),\ \forall\, \psi \in \mathcal{H}, \ \text{ where } \mathrm{d}\mu(z) := \frac{\mathrm{d}^2 z}{\pi}.$$

The function $\psi(z) := \langle\overline{z}|\psi\rangle$ is called the *symbol* of ψ. Using (P4) again, we get

$$\psi(z') \,=\, \int_{\mathbb{C}} K(\overline{z'}, z)\, \psi(z)\, \mathrm{d}\mu(z),$$

where $K(\overline{z'}, z) :=\ \langle z'|z\rangle = \exp\left(-\frac{1}{2}|z'|^2 + \overline{z'}z - \frac{1}{2}|z|^2\right)$ is a *reproducing kernel*:

$$\int_{\mathbb{C}} K(\overline{z'}, z'')\, K(\overline{z''}, z)\ \mathrm{d}\mu(z'') \,=\, K(\overline{z'}, z).$$

From the resolution of the identity (2.5), one gets

$$\langle\psi_1|\psi_2\rangle = \int_{\mathbb{C}} \overline{\psi_1(z)}\,\psi_2(z)\;\mathrm{d}\mu(z)\;,\;\forall\,\psi_1,\,\psi_2\,\in\mathcal{H},$$

that is, the map $W : \mathcal{H} \to L^2(\mathbb{C}, \mathrm{d}\mu)$ defined by $(W\psi)(z) = \psi(z)$ is an *isometry*. Hence $\mathcal{H}_{\mathrm{FB}} := \mathrm{Ran}\, W$ is a closed subspace of $L^2(\mathbb{C}, \mathrm{d}\mu)$. The corresponding projection operator $P_W : L^2(\mathbb{C}, \mathrm{d}\mu) \to \mathcal{H}_{\mathrm{FB}}$ is an *integral* operator with kernel $K(\overline{z'}, z)$. This approach leads to the so-called *Fock-Bargmann representation* of quantum mechanics.

Write

$$\psi(z) = \langle\overline{z}|\psi\rangle = e^{-\frac{1}{2}|z|^2}\, f(z),$$

where f is an *entire* holomorphic function. Then one has the unitary equivalence

$$\mathcal{H} \simeq \mathcal{H}_{\mathrm{FB}} := \left\{ f\,\text{entire},\; \int_{\mathbb{C}} |\,f\,(z)|^2\, e^{-|z|^2}\;\mathrm{d}\mu(z)\; < \infty \right\}.$$

In this Fock-Bargmann realization, one has

$$a = \frac{\mathrm{d}}{\mathrm{d}z}, \quad a^\dagger = \text{ multiplication operator by } z,$$

and the orthonormal basis

$$u_n(z) = \frac{z^n}{\sqrt{n!}}\,, \quad n = 0, 1, 2\ldots.$$

Since $z = 2^{-1/2}\,(q+ip)$, the Fock-Bargmann representation is in fact a *phase space* representation. This fact will prove important for quantization!

2.3. Generalizations

As said above, all four conditions (P1)–(P4) lead to different generalizations of the canonical CS. We list them very briefly.

(i) Minimal uncertainty states (P1):

A state that verifies (P1) is either a CS or a *squeezed state*, that is, a state of the form

$$|\alpha, z\rangle = S(\alpha)|z\rangle,\; S(\alpha) = \exp\left[\tfrac{1}{2}(\alpha {a^\dagger}^2 - \overline{\alpha} a^2)\right],\; \alpha, z \in \mathbb{C}.$$

Such states have been constructed experimentally in quantum optics.

(ii) Eigenvalue property (P2):

This property has led to the construction of *Barut–Girardello CS* for the discrete series representations of $\mathrm{SL}(2,\mathbb{R}) \simeq \mathrm{SU}(1,1) \simeq \mathrm{SO}_o(1,2)$, or, more generally, for any Lie algebra with ladder operators $A, A^\dagger$, for instance, algebras of q-deformed oscillators or supersymmetric systems. In that case the Barut–Girardello CS are defined as eigenvectors of the lowering operator, i.e., solutions of the eigenvalue equation

$$A|\xi\rangle = \xi|\xi\rangle, \quad \xi \in \mathbb{C}.$$

(iii) Group theory (P3):

This approach has led, in successive generality, to

- *Covariant CS* derived from square integrable representations of locally compact groups: $\eta_g = U(g)\eta$.
  Example: 1-D wavelets corresponding to the $ax + b$ group.
- *Gilmore–Perelomov CS* on homogeneous spaces $X_\eta := G/H_\eta$, where H_η is the isotropy subgroup of η up to a phase.
  Examples: the n-D wavelets associated to the similitude group SIM(n), the canonical CS corresponding to $G_{\rm WH}$ or the spin CS stemming from SU(2).
- *Covariant CS* on arbitrary homogeneous spaces.
  Examples: the relativistic CS, corresponding to the Poincaré and Galilei groups.

We will discuss these three types of (generalized) CS in the following sections.

(iv) Reproducing kernel (P4):

This is the essential property and the guiding principle, for instance in the context of quantization. We will touch briefly on that point of view in Section 5.

3. CS derived from group representations

In this section, we shall discuss the three types of CS derived from group representations, in increasing generality. We follow [2] and [3].

3.1. CS from square integrable representations

The framework consists of a locally compact group G, with left Haar measure dg and a strongly continuous UIR U of G on a Hilbert space $\mathcal{H}$. Assume U is *square integrable* (equivalently, U belongs to the discrete series), i.e., there exists a nonzero vector $\eta \in \mathcal{H}$, called an *admissible vector*, such that

$$I(\eta) := \int_G |\langle U(g)\eta|\eta\rangle|^2 \, {\rm d}g < \infty. \tag{3.1}$$

Equivalently, η is admissible if

$$\int_G |\langle U(g)\eta|\phi\rangle|^2 \ {\rm d}g < \infty, \ \forall \phi \in \mathcal{H}. \tag{3.2}$$

Since η is admissible, the vector $\eta_g = U(g)\eta$ is admissible for all $g \in G$, so that the set $\mathcal{A}$ of all admissible vectors is invariant under U. Then, since U is irreducible, either $\mathcal{A} = \{0\}$ or $\mathcal{A}$ is dense in $\mathcal{H}$. In particular, $\mathcal{A} = \mathcal{H}$, i.e., every vector in $\mathcal{H}$ is admissible, if and only if G is unimodular.

Given a fixed admissible vector $\eta \in \mathcal{A}$, the elements of the orbit of η under U are called *covariant CS* and their set is denoted by $\mathcal{S} = \{\eta_g = U(g)\eta, \, g \in G\}$.

Then the map $W_\eta : \mathcal{H} \to L^2(G, \, {\rm d}g)$, defined as

$$(W_\eta\phi)(g) = [c(\eta)]^{-1/2}\langle \eta_g|\phi\rangle, \ \phi \in \mathcal{H}, \ g \in G, \ \text{ where } c(\eta) := I(\eta)/\|\eta\|^2,$$

is called the *coherent state or CS map*. This is the crucial object of the theory, as discovered in the pioneering work of Grossmann *et al.* [21].

The CS map W_η has the following properties.

(1) W_η is an isometry onto a closed subspace $\mathcal{H}_\eta$ of $L^2(G,\, \mathrm{d}g)$: $W_\eta^* W_\eta = I$. Thus $\mathcal{S}$ is a total set : $\mathcal{S}^\perp = \{0\}$. Equivalently, η_g generates a *resolution of the identity*

$$\frac{1}{c(\eta)} \int_G |\eta_g\rangle\langle\eta_g| \,\mathrm{d}g = I. \tag{3.3}$$

(2) The subspace $\mathcal{H}_\eta = W_\eta \mathcal{H} \subset L^2(G,\, \mathrm{d}g)$ is a *reproducing kernel Hilbert space.* The projection operator $P_\eta = W_\eta W_\eta^*$, $P_\eta L^2(G,\, \mathrm{d}g) = \mathcal{H}_\eta$, is an integral operator with reproducing kernel $K_\eta(g,g') = c(\eta)^{-1}\langle\eta_g|\eta_{g'}\rangle$:

$$(P_\eta\Phi)(g) = \int_G K_\eta(g,g')\Phi(g')\,\mathrm{d}g', \quad \Phi \in L^2(G,\, \mathrm{d}g).$$

(3) W_η may be inverted on its range by the adjoint operator, which yields a *reconstruction formula*:

$$\phi = W_\eta^*\Phi = [c(\eta)]^{-1/2} \int_G (W_\eta\phi)(g)\eta_g \,\mathrm{d}g, \quad \Phi = W_\eta\phi \in \mathcal{H}_\eta$$

(4) W_η intertwines U and the left regular representation U_ℓ:

$$W_\eta U(g) = U_\ell(g) W_\eta, \quad \forall\, g \in G$$

Therefore, U is equivalent to a subrepresentation of U_ℓ or, equivalently, U belongs to the discrete series, or W_η is *covariant* under the action of G.

As a matter of fact, square integrable representations behave in many respect like UIRs of compact groups. In particular, they lead to nice orthogonality relations. If G is compact and U is a UIR of G (necessarily finite dimensional), the matrix elements $\langle U(g)\psi|\phi\rangle$ of U satisfy orthogonality relations and, by the Peter-Weyl theorem, they generate an orthonormal basis of $L^2(G,\, \mathrm{d}g)$. This fact underlies many properties of well-known special functions, such as spherical harmonics or Bessel functions. The same result holds for any locally compact group, if U is square integrable. Indeed, if G is a locally compact group and U a square integrable UIR of G, then there exists a unique positive, self-adjoint, invertible operator C in $\mathcal{H}$, with dense domain $D(C)$ equal to $\mathcal{A}$, such that, for η and η' admissible and ϕ, ϕ' arbitrary in $\mathcal{H}$, one has

$$\int_G \overline{\langle\eta'_g|\phi'\rangle}\langle\eta_g|\phi\rangle \,\mathrm{d}g = \langle C\eta|C\eta'\rangle\ \langle\phi'|\phi\rangle.$$

The operator C is called the *Duflo–Moore operator* and is often denoted by $C = K^{-1/2}$. Then, $C = \lambda I,\ \lambda > 0$, if and only if G is unimodular. If G is compact, $C = [\dim \mathcal{H}]^{-1/2} I$. If G is not compact, but unimodular, then $C = [c(\eta)]^{1/2} I$. In that case, $d_U := c(\eta)^{-1}$ is called the formal dimension of U.

Remark: instead of (3.3), one has also the following generalized resolution of the identity, for η, η' admissible such that $\langle C\eta|C\eta'\rangle \neq 0$,

$$\frac{1}{\langle C\eta|C\eta'\rangle} \int_G |\eta'_g\rangle\langle\eta_g| dg = I.$$

3.2. Gilmore–Perelomov CS

It is a fact that the usual admissibility conditions (3.1) or (3.1) are often too strong. This motivates the generalization due to Gilmore [20] and Perelomov [26] (independently).

Given an admissible vector η, let H_η be subgroup of G that leaves it invariant up to a phase (this condition obviously comes from quantum mechanics, where a state is represented, not by a vector, but by a ray in the Hilbert space):

$$U(h)\eta = \exp i\alpha(h)\eta, \ \forall\, h \in H_\eta$$

Therefore, the integrand of the admissibility condition (3.1) depends only on the coset gH_η, not on η itself. Hence we can introduce a weaker admissibility condition on $X_\eta = G/H_\eta$ (we assume that X_η carries an invariant measure ν, a mild restriction):

$$\begin{aligned} c(\eta, \phi) &= \int_{X_\eta} |\langle U(g)\eta|\phi\rangle|^2 \ \mathrm{d}\nu(x) < \infty, \ \ \forall\, \phi \in \mathcal{H} \quad (x := gH_\eta) \\ &= \int_{X_\eta} |\langle U(\sigma(x))\eta|\phi\rangle|^2 \ \mathrm{d}\nu(x) < \infty, \ \ \forall\, \phi \in \mathcal{H}, \end{aligned} \tag{3.4}$$

with an *arbitrary* section $\sigma : X_\eta \to G$ in the principal fibre bundle $\pi : G \to X_\eta$. When these conditions are satisfied, we say that U is *square integrable mod* H_η.

In that case, we get a new set of CS : $\mathcal{S}_\sigma = \{\eta_{\sigma(x)} = U(\sigma(x))\eta, \, x \in X_\eta\}$, called *Gilmore–Perelomov CS* (η has been normalized by $c(\eta, \eta) = 1$). Note that the admissibility condition (3.4) does *not* depend on σ, but $\mathcal{S}_\sigma$ does !

The properties of Gilmore–Perelomov CS are exactly the same as those of the previous CS. First, $\mathcal{S}_\sigma$ is a total set in $\mathcal{H} : \mathcal{S}_\sigma^\perp = \{0\}$. Then, the CS map $W_\eta : \mathcal{H} \to L^2(X_\eta, \mathrm{d}\nu)$, which reads now as $(W_\eta\phi)(x) = \langle \eta_{\sigma(x)}|\phi\rangle$, is an isometry onto a closed subspace $\mathcal{H}_\eta$ of $L^2(X_\eta, \mathrm{d}\nu)$, which leads to the resolution of the identity

$$\int_{X_\eta} |\eta_{\sigma(x)}\rangle\langle\eta_{\sigma(x)}| \, \mathrm{d}\nu(x) = I.$$

Finally, the projection operator $P_\eta = W_\eta W_\eta^* : L^2(X_\eta, \mathrm{d}\nu) \to \mathcal{H}_\eta$ is an integral operator with reproducing kernel $K(x', x) = \langle \eta_{\sigma(x')}|\eta_{\sigma(x)}\rangle$.

As examples of Gilmore–Perelomov CS, we may give:

- If G is the Weyl–Heisenberg group G_{WH}, leading to canonical CS, the CS map W_η is the windowed Fourier or Gabor transform.
- If G is compact, any UIR is square integrable. For instance, if η is a highest weight state, then H_η is the maximal compact subgroup. A typical example is SU(2) with the spin CS.
- For G noncompact semisimple and U square integrable, typical examples are the discrete series representations of SU(1,1) (useful in the description of path integrals) or $\mathrm{Sp}(3, \mathbb{R})$ (nuclear structure theory).

- If G is the similitude group of $\mathbb{R}^n$, $\mathrm{SIM}(n) = \mathbb{R}^n \rtimes (\mathbb{R}_*^+ \times \mathrm{SO}(n))$, then the corresponding CS are the n-dimensional wavelets that we shall discuss in Section 9.1.

3.3. General square integrable covariant CS

In practice, it is often too restrictive to require that the subgroup H in a quotient $X = G/H$ be the invariance subgroup of a given admissible vector, even up to a phase, as Gilmore and Perelomov do. Thus a natural extension of the previous analysis consists in constructing CS over an arbitrary homogeneous space.

Thus we start with a locally compact group G, a strongly continuous UIR U in $\mathcal{H}$ and a closed subgroup H of G. Then we consider the homogeneous space $X := G/H$, equipped with an invariant measure ν (this is a weak restriction). Let $\sigma : X = G/H \to G$ be a Borel section (in the principal bundle $G \to G/H$). Then we say that U is *square integrable mod* (H, σ) for the vector $\eta \in \mathcal{H}$ (or that η is admissible for (U, σ)) if

$$0 < c_X(\eta, \phi) := \int_X |\langle U(\sigma(x))\eta|\phi\rangle|^2 \,\mathrm{d}\nu(x) = \langle \phi | A_\sigma \phi\rangle \ < \infty, \ \ \forall\, \phi \in \mathcal{H},$$

with A_σ a bounded positive *invertible* operator on $\mathcal{H}$, called the frame operator. Note that A_σ^{-1} may be unbounded.

Then coherent states based on X read as $\mathcal{S}_\sigma = \{\eta_{\sigma(x)} := U(\sigma(x))\eta,\, x \in X\}$ (here again we put $c_X(\eta, \eta) = 1$). Clearly these new CS depend on the choice of the section σ. Yet they have essentially the same properties as the previous ones.

(1) $\mathcal{S}_\sigma$ is total (overcomplete) in $\mathcal{H} : \mathcal{S}_\sigma^\perp = \{0\}$.

(2) The range $\mathcal{H}_\eta$ of the linear map (CS map) $W_\eta : \mathcal{H} \to L^2(X, \mathrm{d}\nu)$, given by $(W_\eta \phi)(x) = \langle \eta_{\sigma(x)} | \phi\rangle$, is complete with respect to the new scalar product $\langle \Phi|\Psi\rangle_\eta := \langle \Phi | W_\eta\, A_\sigma^{-1}\, W_\eta^{-1} \Psi\rangle$ and W_η is *unitary* from $\mathcal{H}$ onto $\mathcal{H}_\eta$. The new fact here is the occurence of the operator A_σ^{-1} in the scalar product. Thus we get a resolution of the identity (with weak convergence, as usual):

$$\int_X |\eta_{\sigma(x)}\rangle\langle\eta_{\sigma(x)}| \,\mathrm{d}\nu(x) = A_\sigma .$$

If A_σ^{-1} is bounded, $\mathcal{S}_\sigma$ is a (possibly continuous) *frame*; if $A_\sigma = \lambda I$, $\mathcal{S}_\sigma$ is a *tight frame*.

(3) The orthogonal projection $P_\eta : L^2(X, \mathrm{d}\nu) \to \mathcal{H}_\eta$ is an integral operator K_σ and $\mathcal{H}_\eta$ is a reproducing kernel Hilbert space of functions. The kernel is $K(x', x) = \langle \eta_{\sigma(x')} | A_\sigma^{-1}\, \eta_{\sigma(x)}\rangle$ if $\eta_{\sigma(x)} \in D(A_\sigma^{-1})$, $\forall\, x \in X$; otherwise, the relation must be understood in a distributional sense.

(4) As a consequence, W_η may be inverted on its range by the adjoint operator, $W_\eta^{-1} = W_\eta^*$ on $\mathcal{H}_\eta$, to obtain (again for $\eta_{\sigma(x)} \in D(A_\sigma^{-1}), \forall\, x \in X$)

$$W_\eta^{-1}\Phi = \int_X \ \Phi(x)\, A_\sigma^{-1}\, \eta_{\sigma(x)} \,\mathrm{d}\nu(x), \ \ \Phi \in \mathcal{H}_\eta\, .$$

Examples of this construction are, of course, Gilmore–Perelomov CS, for which H is the stability subgroup of η and $A_\sigma = I$, but also CS for the relativity groups (Euclidean, Galilei, Poincaré), which are *not* of the Gilmore–Perelomov type.

4. Algebraic CS

With retrospect, we realize that the crucial ingredient for the construction of CS is the reproducing kernel, not group theory. Thus we are led to an algebraic formulation, directly generalizing the original definition of Schrödinger (2.2).

4.1. Nonlinear CS

A new class of CS was introduced under the name of *Nonlinear CS*, simply by modifying the standard formula (2.2) for the canonical CS, namely,

$$| z \rangle = e^{-\frac{1}{2}|z|^2} \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}} \, | e_n \rangle, \ z \in \mathbb{C},$$

where $\{e_n\}$ is an orthonormal basis in $\mathcal{H}$. Given an increasing sequence of positive numbers $0 < \varepsilon_1 \leqslant \varepsilon_2 \leqslant \ldots \leqslant \varepsilon_n \leqslant \ldots$, one defines the new CS as

$$|z\rangle = \mathcal{N}(|z|^2)^{-1/2} \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{\varepsilon_n!}} \, |e_n\rangle, \tag{4.1}$$

where $\varepsilon_n! := \varepsilon_1 \varepsilon_2 \ldots \varepsilon_n$ is a generalized factorial. In (4.1), the normalization factor $\mathcal{N}(|z|^2)$ is chosen so that $\langle z|z\rangle = 1$.

These CS are overcomplete and satisfy a resolution of the identity

$$\int_{\mathcal{D}} |z\rangle\langle z| \, \mathcal{N}(|z|^2) \; \mathrm{d}\nu(z, \overline{z}) = I \,,$$

where $\mathcal{D} \subset \mathbb{C}$ is an open disc of radius L, the radius of convergence of $\sum_{n=0}^{\infty} z^n/\sqrt{\varepsilon_n!}$. The measure ν is defined as $\mathrm{d}\nu = \mathrm{d}\theta \, \mathrm{d}\lambda(r)$ (for $z = re^{i\theta}$), where $\mathrm{d}\lambda$ is related to the ε_n through the moment condition:

$$\frac{\varepsilon_n!}{2\pi} = \int_0^L r^{2n} \, \mathrm{d}\lambda(r), \quad n = 0, 1, 2, \ldots .$$

In addition, these CS possess a reproducing kernel, namely,

$$K(\overline{z}, z') = \langle z|z'\rangle = \left[\mathcal{N}(|z|^2)\mathcal{N}(|z'|^2)\right]^{-1/2} \sum_{n=0}^{\infty} \frac{(\overline{z}z')^n}{\varepsilon_n!} \, .$$

4.2. Gazeau-Klauder CS

The so-called Gazeau-Klauder CS are a special case of the general nonlinear CS just described. We consider three cases, depending whether the spectrum of the Hamiltonian is discrete, continuous or mixed.

(1) Discrete spectrum dynamics

Let H be a positive Hamiltonian with purely discrete nondegenerate spectrum $H|e_n\rangle = \omega\varepsilon_n|e_n\rangle$, where $0 = \varepsilon_0 < \varepsilon_1 < \varepsilon_2 < \ldots < \varepsilon_n < \ldots$, is a finite or infinite sequence. The corresponding Gazeau–Klauder or action-angle CS read as

$$|J,\gamma\rangle = \frac{1}{\sqrt{\mathcal{N}(J)}} \sum_{n\geqslant 0} \frac{J^{n/2}}{\sqrt{\varepsilon_n!}} e^{-i\varepsilon_n\gamma}|e_n\rangle, \quad J \geqslant 0,\ \gamma\in\mathbb{R}. \tag{4.2}$$

Of course, we recover the canonical CS if we put $\varepsilon_n = n,\ z = \sqrt{J}\, e^{-in\gamma}$. The CS (4.2) enjoy all the standard properties:

- Continuity of the map $[0,R)\times\mathbb{R} \ni (J,\gamma) \mapsto |J,\gamma\rangle \in \mathcal{H}$, where the convergence radius $R = \overline{\lim}_{n\to\infty} \sqrt[n]{\varepsilon_n!}$ is nonzero and $0 \leqslant J < R$.
- Temporal stability : $e^{-iHt}|J,\gamma\rangle = | J,\gamma+\omega t\rangle$.
- Action identity : $\langle J,\gamma|H|J,\gamma\rangle = \omega J$.
- Resolution of the identity :

$$I = \int |J,\gamma\rangle\langle J,\gamma|\ \mathrm{d}\mu_\rho(J,\gamma) := \lim_{\Gamma\to\infty}\frac{1}{2\Gamma}\int_{-\Gamma}^{\Gamma} d\gamma \int_0^L |J,\gamma\rangle\langle J,\gamma|\ \mathcal{N}(J) w_\rho(J)\ \mathrm{d}J,$$

provided the following moment problem is satisfied :

$$\varepsilon_n! = \int_0^L J^n\, w_\rho(J)\ \mathrm{d}J, \quad w_\rho(J) \geqslant 0.$$

Next we exhibit two explicit examples of Gazeau–Klauder CS.

1. Coulomb-like discrete spectrum:

$$\varepsilon_n = 1 - \frac{1}{(n+1)^2}, \text{ for which } \varepsilon_n! = \frac{1}{2}\frac{n+2}{n+1}\,.$$

The distribution probability $w(J)$ is given by

$$\varepsilon_n! = \int_0^1 J^n w(J)\ \mathrm{d}J \text{ where } w(J) = \frac{1}{J}(1+\delta(J-1^-)).$$

2. Infinite square well and Pöschl–Teller potentials:

Consider the following Hamiltonian for a particle on the line :

$$H = -\frac{1}{2m}\frac{d^2}{dx^2} + V(x) - \frac{\lambda^2}{2m},$$

where

$$V(x) = \begin{cases} \dfrac{4\lambda(\lambda-1)}{\sin^2 x}, & 0 < x < \pi, \\ \infty, & x \leqslant 0,\ x \geqslant \pi, \end{cases}$$

For $\lambda = 1$, this yields the infinite square well and for $\lambda > 1$ the symmetric Pöschl–Teller potential. For these two cases, everything is computable explicitly : $\varepsilon_n = n(n+2\lambda)$; $\omega = \frac{1}{2m}$; $R = \infty$; $w(J)$; the resolution of the identity.

(2) CS for continuum dynamics

Take now a Hamiltonian with positive, nondegenerate, purely continuous spectrum:

$$H|\mathsf{E}\rangle = \mathsf{E}|\mathsf{E}\rangle, \quad 0 \leqslant \mathsf{E} < \overline{\mathsf{E}}.$$

The corresponding normalized CS are

$$|J,\gamma\rangle = \frac{1}{\sqrt{\mathcal{N}_\rho(J)}} \int_0^{\overline{\mathsf{E}}} \frac{J^{\mathsf{E}/2}}{\sqrt{\rho(\mathsf{E})}} e^{-i\gamma\mathsf{E}} |\mathsf{E}\rangle \, \mathrm{d}\mathsf{E},$$

with the normalization condition

$$\mathcal{N}_\rho(J) = \int_0^{\overline{\mathsf{E}}} \frac{J^{\mathsf{E}}}{\rho(\mathsf{E})} \, \mathrm{d}\mathsf{E} < \infty, \text{ for } 0 \leqslant J < R.$$

For these CS, the continuity and temporal stability properties are satisfied and there is a nonnegative probability density $w_\rho(J)$ such that $\int_0^R J^{\mathsf{E}} w_\rho(J) \, \mathrm{d}J = \rho(\mathsf{E})$, from which follows the resolution of the identity

$$I = \int |J,\gamma\rangle\langle J,\gamma| \, \mathrm{d}\mu_\rho(J,\gamma) := \int_{-\infty}^{\infty} d\gamma \int_0^R |J,\gamma\rangle\langle J,\gamma| \, \mathcal{N}_\rho(J) w_\rho(J) \, \mathrm{d}J.$$

(3) CS for discrete and continuum dynamics

For a Hamiltonian having both discrete and continuous spectra, $H = H_D \oplus H_C$, one simply combines the two previous cases.

4.3. More exotic CS

All the coherent states we have encountered so far live in complex Hilbert spaces. However the concept has been extended to a number of more exotic situations. We may mention:

(1) Vector CS over matrix domains

The algebraic definition (4.1) may be generalized to the case where the complex variable z is replaced by a $n \times n$ matrix $\mathcal{Z}$, taking its value in an appropriate domain, chosen in such a way that a resolution of the identity holds true [28]. In addition, the Hilbert space is taken of the form $\mathbb{C}^n \otimes \mathcal{H}$, which leads to the so-called Vector CS (VCS). This approach generalizes canonical CS, including their quaternionic version.

(2) CS over quaternionic Hilbert spaces

A quaternionic Hilbert space is a Hilbert space where the scalars are quaternions instead of complex numbers. Since quaternions are noncommutative, one has to distinguish between left and right quaternionic Hilbert spaces. In such a framework, the whole CS machinery may be developed, following essentially the same procedure as in the case of matrix-valued VCS described above [29].

(3) CS on Hilbert modules

First we may rewrite (4.1) in the following abstract form (this anticipates (5.1) below):

$$|z\rangle = \sum_{n=0}^{\infty} \phi_n(z)\, e_n\,,$$

where $\phi_n \in L^2(\mathcal{D})$ and $\{e_n\}$ is an orthonormal basis in the Hilbert space $\mathcal{K}$. Take now two unital C*-algebras $\mathcal{A}, \mathcal{B}$ and a Hilbert C*-module $\mathcal{E}$ over $\mathcal{B}$, with left action from $\mathcal{A}$ [23]. Next one defines the space $\mathcal{H} = \mathcal{E} \otimes \mathcal{K}$, which is again a Hilbert C*-module under the right action of $\mathcal{B}$. Then one choses an orthonormal basis $\{F_n\}$ in $L^2(X, \mathrm{d}\mu; \mathcal{E})$, the space of square integrable, $\mathcal{E}$-valued functions on the measure space (X, μ) and a basis $\{\phi_n\}$ in the Hilbert space $\mathcal{K}$. Putting all together, one defines the so-called *module-valued CS* as

$$|x, a\rangle = \sum_n aF_n(x) \otimes \phi_n, \ \text{ where } a \in \mathcal{A} \text{ satisfies } aa^* = I_{\mathcal{A}}.$$

These CS $|x, a\rangle \in \mathcal{H}$ then verify all the expected properties, including the resolution of the identity [4]. They also generalize the VCS. A variant of this construction also leads to the Cuntz algebras.

5. Probabilistic aspects, quantization

CS may be defined for a general system in the following way. Take as "observable" space a measure space (X, μ), on which observables are defined as (generalized) functions $f \in L^2(X, \mathrm{d}\mu)$. Assume there is an orthonormal basis $\{\phi_n\}$ in $L^2(X, \mathrm{d}\mu)$ such that $0 < \mathcal{N}(x) := \sum_n |\phi_n(x)|^2 < \infty$ (a.e.). Then consider an abstract Hilbert space $\mathcal{H}$ (the space of quantum states) with orthonormal basis $\{|e_n\rangle\}$ in one-to-one correspondence $|e_n\rangle \leftrightarrow \phi_n$ with the previous basis. In this setup, CS may be defined as

$$X \ni x \mapsto |x\rangle = \frac{1}{\sqrt{\mathcal{N}(x)}} \sum_n \overline{\phi_n(x)} |e_n\rangle \in \mathcal{H}, \tag{5.1}$$

where one imposes

$$\langle x|x\rangle = 1\,, \qquad \int_X |x\rangle\langle x|\, \mathcal{N}(x)\ \mathrm{d}\mu(x) = I_{\mathcal{H}}.$$

Then the set $\{|x\rangle,\, x \in X\}$ may be interpreted as a *frame* on X, that selects a certain point of view on X. Next, the CS (5.1) determine two probability distributions, in Bayesian duality:

- A *prior* distribution on the set of indices, $n \mapsto |\phi_n(x)|^2/\mathcal{N}(x)$.
- A *posterior* distribution on the set of parameters, i.e., the classical measure space (X, μ), $x \mapsto |\phi_n(x)|^2$, parameterized by n.

These probabilistic aspects will play a key role in the quantization procedure, that is, the construction of a quantum system from a classical one. Take again for classical observable a (generalized) function $f(x)$. Typically, X is a phase space and $x = (q, p)$. Then, by *quantization*, we mean a map $f(x) \mapsto A_f$, where A_f is an operator on some (quantum) Hilbert space, satisfying a number of consistency conditions:

- $1 \mapsto I$,
- Poisson bracket $\{f, g\} \mapsto$ commutator $[A_f, A_g]$.

However, this procedure is plagued by a number of problems, such as noncommutativity of operator algebras, domain problems, consistency, ...

A consistent procedure, that avoids these problems, is the so-called *CS or Berezin quantization*, which is defined as follows:

$$f(x) \mapsto A_f := \int_X |x\rangle\langle x|\, f(x)\, \mathcal{N}(x)\, \mathrm{d}\mu(x),$$

where f may be a smooth function, $f \in L^2$ or a distribution. With this definition, $\widehat{A_f} := f(x)$ is called the *contravariant* or the *upper symbol* of A_f and $\check{f}(x) := \langle x|A_f|x\rangle$ the *covariant* or the *lower symbol* of f. For testing the validity of the method, one has to compare $\check{f}(x) \simeq f(x)$, at least in some limit, for instance $\hbar \to 0$, corresponding to the transition quantum $\to$ classical.

For a thorough study of this and other methods of quantization, we refer to the review [1], as well as [19] or [3, Chap. 11]. In particular, the connection between quantization, complex Hermite polynomials and reproducing kernel Hilbert spaces is explored in [5].

6. Remarks on signal processing: Time-frequency representation

The sequel of this paper will be devoted to wavelets, which are in fact a special case of CS, namely those associated to the affine group of the line. However, instead of proceeding along the same lines as in the preceding sections, we will start from scratch, with some general remarks on signal processing. Actually this means that we follow the chronological order in the development of wavelet theory. For more details, we refer to the classical treatises of Daubechies [17] or Mallat [24], and also to our own textbooks [3, 10].

The aim of signal processing is to transform raw signals, as given by receivers of some sort, in such a way that they can be analyzed, compressed and retransmitted. The traditional tool to that effect is simply the Fourier transform

$$s(x) \quad \longleftrightarrow \quad \widehat{s}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\xi x} s(x)\, \mathrm{d}x. \tag{6.1}$$

However, Fourier analysis is not sufficient. On one hand, time localization is totally lost : when does the $\widehat{s}(\xi)$ component occur ? On the other hand, it is very *unstable*: a tiny perturbation of the signal, such as a Dirac δ, perturbs completely the Fourier spectrum. The reason, of course, is that Fourier transform is global.

Therefore, signal analysts turn to *time-frequency* (TF) representations. The idea is that one needs *two* parameters. One, called a, characterizes the frequency,

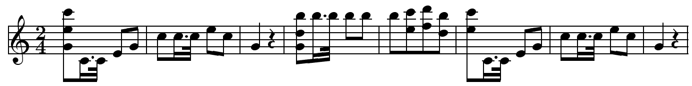


FIGURE 1. A traditional time-frequency representation of a signal (from Mozart's Don Giovanni, Act 1).

the other one, b, indicates the position in the signal. This concept of a TF representation is in fact quite old and familiar. The most obvious example is simply a musical score (Fig. 1)!

If one requires, in addition, the transform to be *linear*, a general TF transform will take the form:

$$s(x) \mapsto S(b,a) = \int_{-\infty}^{\infty} \overline{\psi_{b,a}(x)}\, s(x)\; \mathrm{d}x, \tag{6.2}$$

where s is the signal and $\psi_{b,a}$ the analyzing function (we denote the time variable by x, in view of the extension to higher dimensions). The assumption of linearity is actually nontrivial, for there exists a whole class of quadratic or, more properly, sesquilinear time-frequency representations. The prototype is the so-called Wigner-Ville transform, introduced originally by E.P. Wigner [31] in quantum mechanics (in 1932!) and extended by J. Ville [30] to signal analysis.

Within the class of linear TF transforms, two stand out as particularly simple and efficient, the Windowed or Short-Time Fourier Transform (STFT)[2] and the wavelet transform (WT). For both of them, the analyzing function $\psi_{b,a}$ is obtained by a group action on a basic (or mother) function ψ, only the group differs. The essential difference between the two is in the way the frequency parameter a is introduced. In both cases, b is simply a time translation. The kernels of the two transforms can be written as follows.

- For the *Windowed Fourier Transform*, one chooses

$$\psi_{b,a}(t) = e^{it/a}\, \psi(t-b). \tag{6.3}$$

  Here ψ is a window function and the a-dependence is a modulation ($1/a \sim$ frequency); the window has constant width, but the lower a, the larger the number of oscillations in the window. The associated group is the Weyl-Heisenberg group $\mathrm{G}_{\rm WH}$ Thus, Gabor analysis amounts to an expansion in terms of *canonical CS* !

- For the *wavelet transform*, instead, one takes

$$\psi_{b,a}(t) = \frac{1}{\sqrt{a}}\, \psi\Big(\frac{t-b}{a}\Big). \tag{6.4}$$

  The action of a on the function ψ is a dilation ($a > 1$) or a contraction ($a < 1$). The shape of the function is unchanged, it is simply spread out or squeezed. In particular, the effective support of ψ_{ba} varies as a function of a.

[2]Often called the *Gabor transform*, although Gabor considered only a discretized version [18].

Here the associated group is the $ax + b$ group, the affine group of the line. Thus wavelet analysis follows the pattern of general CS theory.

7. The Continuous Wavelet Transform in 1-D

7.1. Basic formulas, interpretation

As announced above, the basic formulas read, in time domain and in frequency domain, respectively,

$$S(b,a) := \langle \psi_{b,a} | s \rangle = |a|^{-1/2} \int_{-\infty}^{\infty} \overline{\psi\left(a^{-1}(x-b)\right)}\, s(x)\, \mathrm{d}x \tag{7.1}$$

$$= |a|^{1/2} \int_{-\infty}^{\infty} \overline{\widehat{\psi}(a\xi)}\, \widehat{s}(\xi)\, e^{i\xi b}\, \mathrm{d}\xi, \tag{7.2}$$

where $a \neq 0$ is a scale parameter, $b \in \mathbb{R}$ is a translation parameter and the hat denotes a Fourier transform. Thus the pair (b, a) runs over the time-scale half-plane $\mathbb{R}^2_+$. In practice one generally uses the restriction $a > 0$, as in (6.4), which is physically more natural.

In these relations, s is a square integrable function and the analyzing wavelet ψ is assumed to be well localized *both* in the space (or time) domain and in the frequency domain. Here the minimal requirement is that $\psi \in L^1 \cap L^2$, but in practice stronger conditions are usually imposed (like $\psi \in \mathcal{S}$, Schwartz's space of fast decreasing functions). In consequence, the CWT provides good bandpass filtering both in x *and* in ξ.

Moreover, ψ must satisfy the following admissibility condition, which guarantees the invertibility of the WT (see (7.14) below):

$$c_\psi := 2\pi \int_{-\infty}^{\infty} \frac{|\widehat{\psi}(\xi)|^2}{|\xi|}\, \mathrm{d}\xi < \infty. \tag{7.3}$$

In most cases, this condition may be reduced to the requirement that ψ has zero mean (the condition is necessary, and becomes sufficient upon a slight restriction on ψ):

$$\widehat{\psi}(0) = 0 \iff \int_{-\infty}^{\infty} \psi(x)\, \mathrm{d}x = 0. \tag{7.4}$$

In addition, ψ is often required to have a certain number of *vanishing moments:*

$$\int_{-\infty}^{\infty} x^n\, \psi(x)\, \mathrm{d}x = 0,\ n = 0, 1, \ldots N. \tag{7.5}$$

This property improves the efficiency of ψ at detecting singularities in the signal, since it is blind to polynomials up to order N, which constitute the smoother, and less informative, part of a signal (think of a Taylor expansion).

In order to fix ideas, we exhibit here two commonly used, and well documented, wavelets (see Fig. 2).

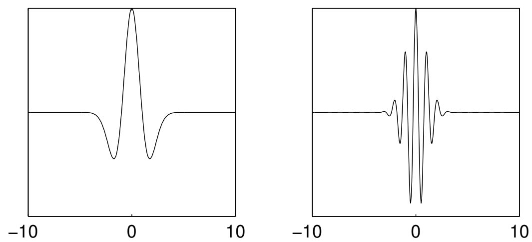


FIGURE 2. (left) The Mexican hat or Marr wavelet; (right) The real part of the Morlet wavelet, for $\xi_o = 5.6$.

1. *The Mexican hat or Marr wavelet*
   This wavelet is simply the second derivative of a Gaussian:
   $$\psi_{\rm H}(x) = (1 - x^2)\, e^{-x^2/2}, \quad \widehat{\psi}_{\rm H}(\xi) = \xi^2\, e^{-\xi^2/2}. \tag{7.6}$$
   It is real and admissible, it has 2 vanishing moments $n = 0, 1$.
2. *The Morlet wavelet*
   This wavelet is essentially a plane wave within a Gaussian window:
   $$\psi_{\rm M}(x) = e^{ik_0 x}\, e^{-x^2/2} + h(x), \quad \widehat{\psi}_{\rm M}(\xi) = e^{-(\xi-\xi_0)^2/2} + \widehat{h}(\xi), \tag{7.7}$$
   Here the correction term h must be added in order to satisfy the admissibility condition (7.4), but in practice one will arrange that this term be numerically negligible ($\leqslant 10^{-4}$) and thus can be omitted (it suffices to choose the basic frequency $|\xi_0|$ large enough, namely, one has to take $|\xi_0| > 5.5$).

These two wavelets have very different properties and, naturally, they will be used in quite different situations. Typically, the Mexican hat is sensitive to singularities in the signal, and it yields a genuine time-scale analysis. On the other hand, since the Morlet wavelet is complex, it will catch the phase of the signal, hence it will be sensitive to frequencies, and will lead to a time-frequency analysis, somewhat closer to a Gabor analysis. In both cases, additional flexibility is obtained by adding a width parameter to the Gaussian.

We must now make more precise the localization conditions on the wavelet ψ. It turns out that, for large scales ($a \gg 1$), the CWT is sensitive to *low frequencies*, and thus yields a rough analysis. On the contrary, for very small scales ($a \ll 1$), the CWT is sensitive to *high frequencies*, that is, small details.

Combining now these localization properties with the fact that the CWT is a convolution with a zero mean filter, we conclude that the CWT provides a local filtering in time (b) and scale (a):

$$S(b,a) \not\approx 0 \quad \Longleftrightarrow \quad \psi_{b,a}(x) \approx s(x).$$

Thus it may be interpreted as a mathematical microscope, with optics ψ, position b, and magnification $1/a$, so that one may consider the CWT as a *singularity detector*. In addition, thanks to the covariance property under scaling (see (7.11) below), the CWT can measure the strength of a singularity, hence it is also a *singularity analyzer*.

7.2. The group-theoretical background

Combining dilations and translations, one gets *affine* transformations of the line

$$y = (b,a)x \equiv ax + b, \quad a \neq 0,\ b \in \mathbb{R},\ x \in \mathbb{R}.$$

Thus $\{(b,a)\} =: G_{\text{aff}} = \mathbb{R} \rtimes \mathbb{R}_*$, the *affine group* of the line.

The action of (b,a) on a signal ψ reads as

$$(U(b,a)\psi)(x) = |a|^{-1/2}\psi\left(\frac{x-b}{a}\right), \tag{7.8}$$

and U is a unitary irreducible representation of G_{aff} in $L^2(\mathbb{R}, \mathrm{d}x)$ (unique up to unitary equivalence). Moreover, U is *square integrable* and a function ψ is admissible if and only if

$$\iint_{G_{\text{aff}}} |\langle U(b,a)\psi|\psi\rangle|^2 \,\frac{\mathrm{d}b\ \mathrm{d}a}{a^2} = c_\psi \,\|\psi\|^2 < \infty, \tag{7.9}$$

where c_ψ is given in (7.3). Restricting to $a > 0$, one gets the *connected* affine group G_{aff}^+ (or $ax+b$ group) and U is a UIR of it in $L^2(\mathbb{R}^+, \mathrm{d}x)$.

7.3. Mathematical properties

Given an admissible wavelet ψ, the CWT $W_\psi : s(x) \mapsto S(b,a)$ is the CS map associated with the representation U of the $ax+b$ group given in (7.8) (with $a > 0$). Hence, it enjoys all properties of CS maps described in the previous sections.

(1) *Covariance* under translation and dilation:

$$W_\psi : s(x - x_0) \mapsto S(b - x_0, a) \tag{7.10}$$

$$W_\psi : \frac{1}{\sqrt{a_0}}\, s\Big(\frac{x}{a_0}\Big) \mapsto S\Big(\frac{b}{a_0}, \frac{a}{a_0}\Big). \tag{7.11}$$

(2) *Energy conservation :*

$$\int_{-\infty}^{\infty} |s(x)|^2 \,\mathrm{d}x = c_\psi^{-1} \iint_{\mathbb{R}^2_+} |S(b,a)|^2 \frac{\mathrm{d}b\,\mathrm{d}a}{a^2}. \tag{7.12}$$

Thus, $|S(b,a)|^2$ may be interpreted as the energy density in the time-scale half-plane $\mathbb{R}^2_+$.

The relation (7.12) means that W_ψ is an *isometry* from the space of signals $L^2(\mathbb{R}, \mathrm{d}x)$ onto a *closed* subspace $\mathcal{H}_\psi$ of $L^2(\mathbb{R}^2_+, \mathrm{d}b\,\mathrm{d}a/a^2)$, namely, the space of wavelet transforms. An equivalent statement is that the wavelet ψ generates a *resolution of the identity* (weak integral, as usual):

$$c_\psi^{-1} \iint_{\mathbb{R}^2_+} |\psi_{b,a}\rangle \,\langle\psi_{b,a}| \frac{\mathrm{d}b\,\mathrm{d}a}{a^2} = I. \tag{7.13}$$

(3) *Reconstruction formula:* As a consequence, W_ψ is invertible on its range $\mathcal{H}_\psi$ by the adjoint map, thus we obtain an exact reconstruction formula:

$$s(x) = c_\psi^{-1} \iint_{\mathbb{R}^2_+} \psi_{b,a}(x)\, S(b,a)\, \frac{\mathrm{d}b\, \mathrm{d}a}{a^2}. \tag{7.14}$$

In other words, we have a representation of the signal $s(x)$ as a linear superposition of wavelets $\psi_{b,a}$ with coefficients $S(b,a)$.

(4) The projection $P_\psi : L^2(\mathbb{R}^2_+,\, \mathrm{d}b\, \mathrm{d}a/a^2) \to \mathcal{H}_\psi$ is an integral operator, with kernel

$$K(b', a'; b, a) = c_\psi^{-1}\, \langle \psi_{b',a'} | \psi_{b,a} \rangle. \tag{7.15}$$

The kernel K is the autocorrelation function of ψ and it is a *reproducing kernel*. Indeed, the function $f \in L^2(\mathbb{R}^2_+,\, \mathrm{d}b\, \mathrm{d}a/a^2)$ is the WT of a certain signal if and only if it satisfies the *reproduction property*

$$f(b', a) = c_\psi^{-1} \iint_{\mathbb{R}^2_+} \langle \psi_{b',a'} | \psi_{b,a} \rangle\, f(b,a)\, \frac{\mathrm{d}b\, \mathrm{d}a}{a^2}\,. \tag{7.16}$$

This means that the CWT is a highly redundant representation!

The last statement implies that the full information must be contained is a small subset of the half-plane. This property may be exploited in two ways: either one takes a discrete subset, which leads to the theory of *frames*, or one considers only the lines of local maxima, called *ridges*.

8. Discrete wavelet transforms in 1-D

8.1. Discretization of the CWT

Of course, the CWT must be discretized for numerical implementation. This raises the question of choosing an adequate sampling grid. Typically, one says that a discrete lattice $\Gamma = \{b_{jk}, a_j, j, k \in \mathbb{Z}\} \subset \mathbb{R}^2_+$ yields a good discretization if the following *exact* relation holds:

$$s = \sum_{j,k \in \mathbb{Z}} \langle \psi_{jk}, s \rangle \widetilde{\psi}_{jk}, \tag{8.1}$$

with $\psi_{jk} := \psi_{b_{jk}, a_j}$ and $\widetilde{\psi}_{jk}$ is a function explicitly constructible from ψ_{jk}. However, this approach in general leads to *frames*, not bases. Let us recall the basic definition (adapted to the present case). A frame in the Hilbert space $\mathcal{H}$ is a family $\{\psi_{jk}\}$ of vectors for which there exist two constants $\mathsf{m} > 0, \mathsf{M} < \infty$ such that

$$\mathsf{m}\, \|s\|^2 \;\leqslant\; \sum_{j,k\, \in\, \mathbb{Z}} |\langle \psi_{jk} | s \rangle|^2 \;\leqslant\; \mathsf{M}\, \|s\|^2,\ \forall\, s \in \mathcal{H}.$$

The upper bound means that the map $s \mapsto \{\langle \psi_{jk} | s \rangle\}$ is continuous from $\mathcal{H}$ to l^2, whereas the lower bound guarantees the numerical stability of the inverse map [17]. The constants m, M are called the frame bounds. If $\mathsf{m} = \mathsf{M} \neq 1$, the frame is said to be tight. Clearly, if $\mathsf{m} = \mathsf{M} = 1$ and $\|\psi_{jk}\| = 1$, for all $j, k \in \mathbb{Z}$, this reduces to an orthonormal basis.

The main question now is, given a wavelet ψ, to find a lattice Γ such that $\{\psi_{jk}\}$ is a good frame. A detailed analysis then shows that the expansion (8.1) converges essentially as a power series in the quantity $|\mathsf{M}/\mathsf{m} - 1|$, thus a good lattice must satisfy $|\mathsf{M}/\mathsf{m} - 1| \ll 1$. The key to the solution is to choose a lattice adapted to the geometry, for instance the *dyadic* grid $a_j = 2^{-j}$, $b_{jk} = k \cdot 2^{-j}$, which leads to

$$\psi_{jk}(x) = 2^{j/2}\psi(2^j\, x - k), \quad j, k \in \mathbb{Z}.$$

For example, both the Mexican hat and the Morlet wavelets give good, but non tight frames.

8.2. The discrete WT (DWT)

For many applications, such as the description and analysis of signals, a tight frame is as good as an orthonormal basis. Yet, a tight frame is still redundant, which implies that the coefficients are not unique, nor statistically independent. Therefore, when it comes to managing large amounts of data and compressing them, an orthonormal basis becomes mandatory, and this requires a different approach (almost orthogonal to the previous one).

One of the successes of the WT was the discovery that it is possible to construct functions ψ for which $\{\psi_{jk}, j, k \in \mathbb{Z}\}$ is indeed an orthonormal basis of $L^2(\mathbb{R})$, while keeping the good properties of wavelets, including space *and* frequency localization. In addition, this approach yields fast algorithms, and this is the key to the usefulness of wavelets in many applications.

The construction is based on two facts. First, almost all examples of orthonormal bases of wavelets can be derived from a multiresolution analysis, and then the whole construction may be transcripted into the language of Quadrature Mirror Filters (QMF), familiar in the signal processing literature.

A *multiresolution analysis* of $L^2(\mathbb{R})$ is an increasing sequence of closed subspaces

$$\ldots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \ldots$$

with $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$ and $\bigcup_{j\in\mathbb{Z}} V_j$ dense in $L^2(\mathbb{R})$, and such that

(1) $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}$;
(2) There exists a function $\phi \in V_0$, called a *scaling function*, such that the family $\{\phi(x-k), k \in \mathbb{Z}\}$ is an orthonormal basis of V_0.

Combining conditions (1) and (2), one sees that $\{\phi_{jk}(x) := 2^{j/2}\phi(2^j x - k),\ k \in \mathbb{Z}\}$ is an orthonormal basis of V_j. The space V_j can be interpreted as an *approximation* space at resolution 2^j. Defining now

$$V_j \oplus W_j = V_{j+1}, \tag{8.2}$$

we see that W_j contains the additional *details* needed to improve the resolution from 2^j to 2^{j+1}. Thus one gets the decomposition

$$L^2(\mathbb{R}) = \bigoplus_{j\in\mathbb{Z}} W_j = V_{j_0} \oplus \left(\bigoplus_{j=j_0}^{\infty} W_j \right), \tag{8.3}$$

if one chooses a lowest resolution level j_0. In practice, one starts from a sampled signal, taken in some V_J, and then the decomposition (8.3)) is replaced by the finite representation

$$V_J = V_{j_o} \oplus \left(\bigoplus_{j=j_0}^{J-1} W_j \right). \tag{8.4}$$

The crucial theorem then asserts the existence of a function ψ, sometimes called the *mother wavelet*, explicitly computable from ϕ, such that $\{\psi_{jk}(x) := 2^{j/2}\psi(2^j x - k), k \in \mathbb{Z}\}$ is an orthonormal basis of W_j. Then $\{\psi_{jk}, j, k \in \mathbb{Z}\}$ constitutes an orthonormal basis of $L^2(\mathbb{R})$: these are the *orthonormal wavelets*. Well-known examples include the Haar wavelets, the B-splines, and the various Daubechies wavelets.

Now a natural question is to decide which version should be used in a concrete case, the (discretized) CWT or the DWT? As usual in a wavelet context, the answer depends on the application at hand:

- For feature detection, the CWT is preferable, since no *a priori* choice is made for a, b; the CWT is more flexible and also more robust to noise, but only frames will be available in general.
- For large amount of data or data compression, one should use the DWT; it yields orthonormal bases, it is faster, but it is also more rigid.

This last aspect explains why a number of generalizations have been introduced in order to alleviate these drawbacks of the DWT, such as biorthogonal wavelets, wavelet packets, continuous wavelet packets (integrated wavelets), redundant WT (on a rectangular lattice) or "Second generation" wavelets (the so-called lifting scheme). For all these, we refer to the textbooks [10, 17, 24].

9. Wavelet analysis of 2-D images

9.1. Basic formulas and interpretation

Now we turn to the two-dimensional CWT, which has become a major tool in image processing. In this context, an image is a two-dimensional signal of finite energy, represented by a complex-valued function $s \in L^2(\mathbb{R}^2, \mathrm{d}^2\vec{x})$. Given an image, all the geometric operations we want to apply to it are obtained by combining three elementary transformations, namely, rigid translations in the plane of the image, dilations or scaling (global zooming in and out) and rotations. Explicitly, the transformations act on $\vec{x} \in \mathbb{R}^2$ in the familiar way:

$$
\begin{array}{rl}
\text{(i)} & \text{translation by } \vec{b} \in \mathbb{R}^2 : \vec{x} \mapsto \vec{x}\,' = \vec{x} + \vec{b}, \\
\text{(ii)} & \text{dilation by a factor } a > 0 : \vec{x} \mapsto \vec{x}\,' = a\vec{x}, \\
\text{(iii)} & \text{rotation by an angle } \theta : \vec{x} \mapsto \vec{x}\,' = r_\theta(\vec{x}),
\end{array}
$$

where

$$
r_\theta \equiv \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \; 0 \leqslant \theta < 2\pi,
$$

is the familiar 2×2 rotation matrix.

Then dilations, translations, and rotations together constitute the *similitude group* of the plane $\mathrm{SIM}(2) = \mathbb{R}^2 \rtimes (\mathbb{R}^+_* \times \mathrm{SO}(2))$, with action $\vec{y} = (\vec{b}, a, \theta)\vec{x} = a r_\theta \vec{x} + \vec{b}$. This transformation is realized by the following unitary map on finite energy signals:

$$
\big[U(\vec{b}, a, \theta)s\big](\vec{x}) = s_{\vec{b},a,\theta}(\vec{x}) := a^{-1} s(a^{-1}\, r_{-\theta}(\vec{x} - \vec{b})) \tag{9.1}
$$

and U is a unitary irreducible representation of SIM(2) in $L^2(\mathbb{R}^2,\, \mathrm{d}^2\vec{x})$, unique up to unitary equivalence. In addition, U is *square integrable* and a function ψ is admissible if and only if

$$
\iiint_{\mathrm{SIM}(2)} \left|\langle U(\vec{b}, a, \theta)\psi|\psi\rangle\right|^2 \; \mathrm{d}^2\vec{b}\, \frac{\mathrm{d}a}{a^3}\, \mathrm{d}\theta \; < \infty, \tag{9.2}
$$

where $a^{-3}\, \mathrm{d}^2\vec{b}\, \mathrm{d}a\, \mathrm{d}\theta$ is the left Haar measure on SIM(2).

In terms of this action, the basic formulas for the 2-D CWT read

$$
\begin{aligned}
S(\vec{b}, a, \theta) := \langle \psi_{\vec{b},a,\theta}|s\rangle &= a^{-1} \int_{\mathbb{R}^2} \overline{\psi(a^{-1}\, r_{-\theta}(\vec{x} - \vec{b}))}\, s(\vec{x}) \; \mathrm{d}^2\vec{x} && (9.3) \\
&= a \int_{\mathbb{R}^2} e^{i\vec{b}\cdot\vec{k}}\, \overline{\widehat{\psi}(a r_{-\theta}(\vec{k}))}\, \widehat{s}(\vec{k}) \; \mathrm{d}^2\vec{k}. && (9.4)
\end{aligned}
$$

As in 1-D, besides the usual conditions of localization, we have to impose an *admissibility* condition on the wavelet ψ, namely,

$$
c_\psi := (2\pi)^2 \int_{\mathbb{R}^2} \frac{|\widehat{\psi}(\vec{k})|^2}{|\vec{k}|^2} \; \mathrm{d}^2\vec{k} < \infty, \tag{9.5}
$$

which again may be replaced in practice by the following necessary condition:

$$
\widehat{\psi}(\vec{0}) = 0 \iff \int_{\mathbb{R}^2} \psi(\vec{x})\, \mathrm{d}^2\vec{x} = 0. \tag{9.6}
$$

Besides the previous requirements, one can improve the efficiency of the wavelets by imposing additional properties, such as restrictions on the support of ψ and of $\widehat{\psi}$, or *vanishing moments*, up to order $N \geqslant 1$ ($N = 0$ yields the admissibility condition) :

$$
\int_{\mathbb{R}^2} x^\alpha\, y^\beta\, \psi(\vec{x}) \; \mathrm{d}^2\vec{x} \; = 0, \quad \vec{x} = (x, y), \quad 0 \leqslant \alpha + \beta \leqslant N.
$$

This will improve the efficiency at detecting singularities in the signal, since the transform is then blind to smoothest part of the signal, i.e., a polynomial of degree up to N, which contains little information, in general. For instance, if $N = 1$, the transform erases any linear *trend* in the signal, such as a linear gradient of luminosity.

Altogether, the mathematical properties of the 2-D CWT are essentially the same as in the 1-D case, so we list them without further comment.

(1) *Energy conservation*

$$c_\psi^{-1} \iiint_{\mathrm{SIM}(2)} |S(\vec{b}, a, \theta)|^2 \; \mathrm{d}^2\vec{b} \, \frac{\mathrm{d}a}{a^3} \, \mathrm{d}\theta = \int_{\mathbb{R}^2} |s(\vec{x})|^2 \; \mathrm{d}^2\vec{x}, \tag{9.7}$$

i.e., the 2-D CWT is an isometry from the space of signals $L^2(\mathbb{R}^2)$ onto a closed subspace of $L^2(\mathrm{SIM}(2))$, the space of wavelet transforms.

(2) *Reconstruction formula*

Inverting the CWT by the adjoint map, we get

$$s(\vec{x}) = c_\psi^{-1} \iiint_{\mathrm{SIM}(2)} \psi_{\vec{b},a,\theta}(\vec{x}) \, S(\vec{b}, a, \theta) \; \mathrm{d}^2\vec{b} \, \frac{\mathrm{d}a}{a^3} \, \mathrm{d}\theta, \tag{9.8}$$

i.e., we have a decomposition of the signal in terms of the analyzing wavelets $\psi_{\vec{b},a,\theta}$, with coefficients $S(\vec{b}, a, \theta)$.

(3) *Reproduction property (reproducing kernel)*

$$S(\vec{b}', a', \theta') = c_\psi^{-1} \iiint_{\mathrm{SIM}(2)} \langle \psi_{\vec{b}',a',\theta'} | \psi_{\vec{b},a,\theta} \rangle \, S(\vec{b}, a, \theta) \, \mathrm{d}^2\vec{b} \, \frac{\mathrm{d}a}{a^3} \, \mathrm{d}\theta. \tag{9.9}$$

(4) *Covariance*

The CWT is *covariant* under translations, dilations and rotations, that is, under SIM(2): the correspondence $W_\psi : s(\vec{x}) \mapsto S(\vec{b}, a, \theta)$ implies the following ones

$$\begin{array}{rcl} s(\vec{x} - \vec{b}_o) & \mapsto & S(\vec{b} - \vec{b}_o, a, \theta), \\ a_o^{-1} s(a_o^{-1}\vec{x}) & \mapsto & S(a_o^{-1}\vec{b}, a_o^{-1} a \theta), \\ s(r_{\theta_o}(\vec{x})) & \mapsto & S(r_{-\theta_o}(\vec{b}), a, \theta - \theta_o). \end{array}$$

Note that translation covariance ("shift invariance") is lost in the standard formulation of the discrete WT, based on multiresolution, which creates problems in pattern recognition, for instance.

Since all the formulas are almost identical in 1-D and in 2-D, the interpretation of the CWT as a singularity analyzer still holds. In particular, the analysis is most efficient at high spatial frequencies or small scales, thus it provides a good detection of *discontinuities* in images, such as point singularities (contours, corners) or directional features (edges, segments, ...). Hence the 2-D CWT may be seen as a mathematical directional microscope with optics ψ, global magnification $1/a$, and orientation tuning parameter θ; it is a detector and analyzer of singularities.

As a matter of fact, the same analysis may be performed almost *verbatim* for n-dimensional signals and wavelets. The group is now the similitude group of $\mathbb{R}^n$, $\mathrm{SIM}(n) = \mathbb{R}^n \rtimes (\mathbb{R}^+_* \times \mathrm{SO}(n))$, and the CWT is constructed in the same way as for $n = 2$, all formulas are basically the same. For instance, the admissibility condition (9.5) is essentially the same, the only difference is that the factor $|\vec{k}|^2$ in the denominator is replaced by $|\vec{k}|^n$. Therefore, the interpretation of this n-dimensional CWT will also remain the same. The case $n = 3$ may find applications in fluid dynamics, for example.

9.2. Choice of the analyzing wavelet

Even more than in 1-D, it is important to choose a wavelet that is well adapted to the problem at hand. One can distinguish two classes, isotropic wavelets and direction sensitive wavelets.

(i) *Isotropic wavelets*

If one decides to perform a pointwise analysis, or if directions are irrelevant, it is sufficient to use a rotation invariant wavelet. Two standard examples are:

1. *The 2-D Mexican hat wavelet*

   This wavelet (originally introduced by Marr in his pioneering work on vision [25]) is simply the Laplacian of a Gaussian :

$$\psi_{\mathrm{H}}(\vec{x}) = (2 - |\vec{x}|^2)\, \exp(-\tfrac{1}{2}|\vec{x}|^2), \quad \widehat{\psi_{\mathrm{H}}}(\vec{k}) = |\vec{k}|^2\, \exp(-\tfrac{1}{2}|\vec{k}|^2). \tag{9.10}$$

2. *The Difference-of-Gaussians or DOG wavelet*

$$\psi_{\mathrm{D}}(\vec{x}) = \tfrac{1}{2\alpha^2}\, \exp(-\tfrac{1}{2\alpha^2}|\vec{x}|^2) - \exp(-\tfrac{1}{2}|\vec{x}|^2) \qquad (0 < \alpha < 1). \tag{9.11}$$

   This is a very good substitute to the Mexican hat, for $\alpha^{-1} = 1.6$, they are almost undistinguishable. It is widely used in psychophysics.

Here again, in both cases, the efficiency may be improved by adding a width parameter in the Gaussian.

(ii) *Directional wavelets*

On the other hand, if the goal is to detect directional features in an image, or to perform directional filtering, clearly one should resort to a direction sensitive wavelet. The most efficient solution is a *directional* wavelet, that is, a wavelet $\psi(\vec{x})$ such that the numerical support of its Fourier transform $\widehat{\psi}(\vec{k})$ is contained in a convex cone with apex at the origin. Two useful examples are as follows.

1. *The 2-D Morlet wavelet*

$$\widehat{\psi}_{\mathrm{M}}(\vec{k}) = \sqrt{\epsilon}\, \Big(\exp(-\tfrac{1}{2}|A^{-1}(\vec{k} - \vec{k}_0)|^2) - h(\vec{k}) \Big), \tag{9.12}$$

   where $A = \mathrm{diag}[\epsilon^{-1/2}, 1], \epsilon \geqslant 1$, is a 2×2 anisotropy matrix and, as in 1-D, the correction term is required in order to satisfy the admissibility condition, and here too, it is negligible for $|\vec{k}_0| \geqslant 5.6$. The Morlet wavelet is directional, but has poor aperture selectivity. It is shown in Fig. 3 for $\epsilon = 2, \vec{k}_0 = (0, 6)$.

2. *Conical wavelets*, with support in the convex cone

$$\mathcal{C}(-\alpha,\alpha) := \{\vec{k} \in \mathbb{R}^2 : -\alpha \leqslant \arg \vec{k} \leqslant \alpha,\ \alpha < \pi/2\}$$

A very useful example is the Gaussian conical wavelet:

$$\widehat{\psi}^{\mathrm{GC}}_{lm} = \begin{cases} (\vec{k}\cdot\vec{e}_{-\widetilde{\alpha}})^l (\vec{k}\cdot\vec{e}_{\widetilde{\alpha}})^m \ \exp(-\frac{1}{2}k_x^2),\ \vec{k} \in \mathcal{C}(-\alpha,\alpha) \\ 0,\ \text{otherwise} \end{cases} \tag{9.13}$$

where $\vec{e}_\phi$ is the unit vector in the direction ϕ and $\widetilde{\alpha} = -\alpha + \pi/2$, so that $\vec{e}_{-\alpha}\cdot\vec{e}_{\widetilde{\alpha}} = \vec{e}_{\alpha}\cdot\vec{e}_{-\widetilde{\alpha}} = 0$. The first two factors play the role of vanishing moments on the boundaries of the cone. The frequency space version of this so-called Gaussian conical wavelet is shown in Fig. 3, in the case $l = m = 4$, $\alpha = 10°$. This wavelet will be used in Section 11 for motion estimation.

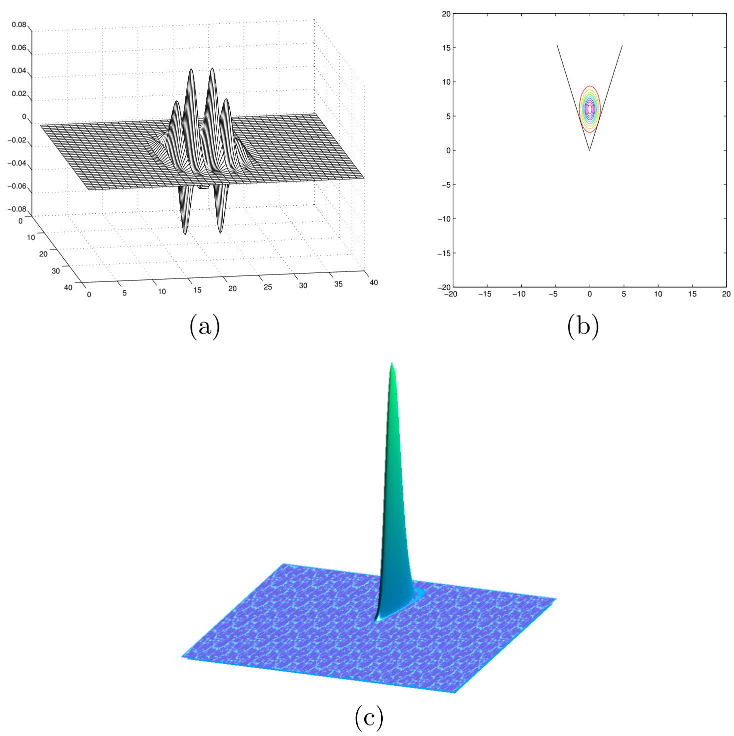


FIGURE 3. Two directional wavelets: The 2-D Morlet wavelet, (a) in position space and (b) in spatial frequency space; (c) The Gaussian conical wavelet, in spatial frequency space.

10. Extending the CWT to curved manifolds

10.1. Generalities

Many situations in physics yield data on non-flat manifolds, for instance:

- A *sphere:* geophysics, cosmology (CMB), statistics, ...;
- A *two-sheeted hyperboloid:* cosmology (an open expanding model of the universe), optics (catadioptric image processing, where a sensor overlooks a hyperbolic mirror);
- A *paraboloid:* optics (catadioptric image processing on a parabolic mirror);
- A *torus:* nuclear fusion (Tokamak), loop quantum gravity.

How can one find suitable analysis tools?

A possible solution is to extend the *continuous* wavelet transform. Translation of the wavelet may be achieved by an isometry of the manifold, i.e., an element of SO(3) or SO(1,2), for the two first cases above (although this does not work always, see below). As for dilations *on* the manifold, they have to be defined in each case. This being done, dilation controls locality of the CWT. However, in practice, the usual CWT works with discrete frames, hence one needs also discrete wavelet frames on the manifold. An alternative solution is to extend the *discrete* wavelet transform.

Let us look first at the general problem. Given a manifold $\mathcal{M}$, how to derive a CWT on it? First, one should identify the operations to be performed on the finite energy signals living on $\mathcal{M}$, i.e., functions in $L^2(\mathcal{M}, \mathrm{d}\mu)$, with μ a suitable measure on $\mathcal{M}$, namely,

(i) Motions are provided by the isometries of $\mathcal{M}$;
(ii) Dilations *on* $\mathcal{M}$ (zoom in/out) should form a one-parameter abelian group $A \sim \mathbb{R}_+$.

If $\mathcal{M}$ admits a global isometry group $G_{\rm iso}$, merge it with the dilation group A into a global group G. Next, find a unitary irreducible, square integrable representation U of G in $L^2(\mathcal{M}, \mathrm{d}\mu)$ and write (we assume that μ is G-invariant)

$$\psi_g(\zeta) := [U(g)\psi](\zeta) = \psi(g^{-1}\zeta), \; g \in G, \zeta \in \mathcal{M}.$$

Then a function $\psi \in L^2(\mathcal{M}, \mathrm{d}\mu)$ is an admissible wavelet if

$$c_\psi := \|\psi\|^{-2} \int_G |\langle U(g)\psi|\psi\rangle|^2 \; \mathrm{d}\mu_\ell(g) < \infty, \tag{10.1}$$

where μ_ℓ is the left Haar measure on G.

Then the CWT of $f \in L^2(\mathcal{M}, \mathrm{d}\mu)$ with respect to the admissible wavelet ψ is defined as

$$W_\psi f(g) := \langle \psi_g | f\rangle = \int_{\mathcal{M}} \overline{\psi(g^{-1}\zeta)}\, f(\zeta)\, \mathrm{d}\mu(\zeta), \; g \in G\,.$$

The corresponding *reconstruction formula* is then given by

$$f(\zeta) = c(\psi)^{-1/2} \int_G W_\psi f(g)\, \psi_g(\zeta)\, \mathrm{d}\mu_\ell(g)\,.$$

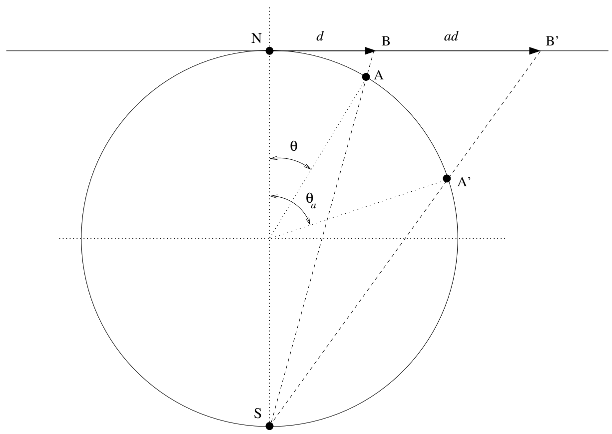


FIGURE 4. The stereographic dilation D_a : A $\mapsto$ A' around the North Pole.

Note, however, that the method does not always work. Some manifolds (e.g. a paraboloid) don't have a global isometry group and, even if the latter exists, the resulting global group G need not have a square integrable UIR.

10.2. The WT on the two-sphere

This is the approach we have followed for the 1-D and the 2-D CWT. Now we illustrate the procedure by constructing a CWT on the two-sphere [7, 8, 9, 13], following the scheme developed in Section 3.3.

As usual, the origin of the spherical CWT consists in the affine transformations on $\mathbb{S}^2$, namely, motions, realized by rotations $\varrho \in \mathrm{SO}(3)$, and dilations. The problem is that they do *not* commute ! Moreover, one cannot build a semidirect product of SO(3) and $\mathbb{R}^+_*$, the only extension of SO(3) by $\mathbb{R}^+_*$ is their *direct product*. A way out of this dilemma is to embed the two factors into the Lorentz group $\mathrm{SO}_o(3,1)$, by the Iwasawa decomposition

$$\mathrm{SO}_o(3,1) = \mathrm{SO}(3) \cdot A \cdot N,$$

where $A \sim \mathrm{SO}_o(1,1) \sim \mathbb{R} \sim \mathbb{R}^+_*$ consists in boosts in the z-direction and $N \sim \mathbb{C}$. The justification of this move is that the Lorentz group $\mathrm{SO}_o(3,1)$ is the *conformal* group both of the sphere $\mathbb{S}^2$ and of the tangent plane $\mathbb{R}^2$, say, at the North Pole.

The Lorentz group $\mathrm{SO}_o(3,1)$ acts transitively on $\mathbb{S}^2$. Then an explicit computation (with help of the Iwasawa decomposition) shows that a boost in z-direction corresponds to a pure dilation, namely, a *stereographic dilation* D_a, as shown in Fig. 4. The dilation proceeds in three steps: (i) Project the initial point A stereographically onto the plane tangent at the North Pole N and get B ; (ii) Dilate B radially around N by a factor a, to B' ; (iii) Project back B' to the sphere and get A' . Then D_a : A $\mapsto$ A' is the required dilation on the sphere.

The Lorentz group $\mathrm{SO}_o(3,1)$ has a natural UIR in the Hilbert space $L^2(\mathbb{S}^2,\, \mathrm{d}\mu)$, where $\mathrm{d}\mu := \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi$ is the usual measure on the sphere, viz.

$$[U(g)f]\,(\omega) = \lambda(g,\omega)^{1/2}\, f\left(g^{-1}\omega\right),\; g \in \mathrm{SO}_o(3,1),\; f \in L^2(\mathbb{S}^2,\, \mathrm{d}\mu),$$

where $\lambda(g,\omega)$ is a Radon-Nikodym derivative, required by the noninvariance of the measure μ under dilation.

Now the parameter space of the spherical CWT is $X := \mathrm{SO}_o(3,1)/N \simeq \mathrm{SO}(3)\cdot\mathbb{R}^+_*$. Thus we introduce a section $\sigma : X \to \mathrm{SO}_o(3,1)$ and consider the reduced representation $U(\sigma(\varrho,a))$. The natural (Iwasawa) section is $\sigma_I(\varrho,a) = \varrho\, a,\; \varrho \in \mathrm{SO}(3),\, a \in A$, and it yields

$$U(\sigma_I(\varrho,a)) = U(\varrho\, a) = U(\varrho)U(a) = R_\varrho\, D_a,$$

where R_ϱ is the unitary operator representing $\varrho \in \mathrm{SO}(3)$.

The UIR U is square integrable on X, that is, there exists nonzero (admissible) vectors $\psi \in L^2(\mathbb{S}^2,\, \mathrm{d}\mu)$ such that [$\mathrm{d}\varrho$ is the left Haar measure on SO(3)]

$$\int_X |\langle U(\sigma_I(\varrho,a))\psi|\phi\rangle|^2 \;\mathrm{d}\varrho\, \frac{\mathrm{d}a}{a^3} := \langle\phi|A_\psi\phi\rangle < \infty,\; \forall\, \phi \in L^2(\mathbb{S}^2,\, \mathrm{d}\mu)\,,$$

where the frame operator A_ψ is bounded, positive and invertible. Actually A_ψ is a Fourier multiplier, i.e., it is a multiplication operator in the Fourier representation. Then, given an admissible vector ψ, the corresponding CWT reads, with $U(\sigma_I(\varrho,a)) = R_\varrho\, D_a$:

$$\begin{aligned} F_\psi(\varrho,a) &:= \langle U(\sigma_I(\varrho,a))\psi|f\rangle \\ &= \int_{\mathbb{S}^2} \overline{[R_\varrho\, D_a\psi](\zeta)}\; f(\zeta)\, \mathrm{d}\mu(\zeta), \quad f \in L^2(\mathbb{S}^2,\, \mathrm{d}\mu). \end{aligned} \tag{10.2}$$

Moreover, given any admissible axisymmetric wavelet ψ, the family $\{\psi_{\varrho,a} := R_\varrho\, D_a\psi,\, (\varrho,a) \in X\}$ is a continuous frame, that is, there exist two constants $\mathsf{m} > 0$ and $\mathsf{M} < \infty$ such that

$$\mathsf{m}\, \|\phi\|^2 \;\leqslant \int_X |\langle\psi_{\varrho,a}|\phi\rangle|^2 \;\mathrm{d}\varrho\, \frac{\mathrm{d}a}{a^3} \;\leqslant \mathsf{M}\, \|\phi\|^2,\; \forall\, \phi \in L^2(\mathbb{S}^2,\, \mathrm{d}\mu).$$

The relation (10.2) yields a full spherical CWT, with correct Euclidean limit, i.e., the spherical CWT on a sphere of radius R tends to the usual 2-D CWT in the tangent plane when $R \to \infty$ (here one uses the technique of group contraction for evaluating the limit $X = \mathrm{SO}(3)\cdot\mathbb{R}^+_* \to \mathrm{SIM}(2)$ and similarly for the corresponding representations). In addition, one may construct spherical frames, both semi-continuous (only the scale is discrete) and fully discrete, but one needs a generalized notion of frame (weighted, controlled frames). In addition, the method extends to the n-sphere.

As for the spherical DWT, there are many methods in the literature on approximation theory. An alternative consists in lifting the plane 2-D DWT from the tangent plane at the North Pole onto $\mathbb{S}^2$ by inverse stereographic projection. In that way, one obtains locally supported orthonormal wavelet bases on $\mathbb{S}^2$ [27].

10.3. The CWT on other manifolds

The techniques developed for the two-sphere may be generalized to other manifolds, in particular, the so-called *conic sections* (the sphere, the two-sheeted hyperboloid, the paraboloid). In all cases, a key ingredient for designing a dilation *on* the manifold is a suitable projection on a simpler one (a tangent cone, a tangent plane) [11, 12]. A detailed description may be found in [3, Chap.15].

1. *The two-sheeted hyperboloid*

   Since the two-sheeted hyperboloid $\mathbb{H}^2$ is the dual of the two-sphere (in the sense of differential geometry), it is not surprising that a group-theoretical construction similar to the one on $\mathbb{S}^2$ works well, using the isometry group SO(2,1) instead of SO(3). Thus one gets a fully developed CWT, but no results about frames are known [14].

2. *The paraboloid*

   The paraboloid $\mathbb{P}^2$ is a singular case, which can be obtained from both $\mathbb{S}^2$ and $\mathbb{H}^2$ by a limiting procedure. However, no (large) isometry group is available, so that the previous method does not work. An alternative construction is possible by lifting the CWT (or the DWT) vertically from the tangent plane onto the paraboloid.

3. *The two-torus*

   The two-torus $\mathbb{T}^2$ is isomorphic to the product of two circles, $\mathbb{T}^2 \simeq \mathbb{S}^1 \times \mathbb{S}^1 \simeq \mathrm{SO}(2) \times \mathrm{SO}(2)$. This fact suggests to exploit the WT on circle, which is analogous to that on the two-sphere, by the same construction, replacing SO(3,1) by SO(2,1). Now, since SO(2,2) $\simeq \mathrm{SO}(2,1) \times \mathrm{SO}(2,1)/\mathbb{Z}_2$, one may go one step further and use SO(2,2). However, this will introduce *two* independent dilations instead of a single, global one. This difficulty may then circumvented by combining SO(2,2) with the modular group $\mathrm{SL}(2,\mathbb{Z})$. In that way one may obtain both a CWT and discrete wavelet frames on $\mathbb{T}^2$ [16].

4. *More general manifolds*

   In the case of a general (two-dimensional) manifold $\mathcal{M}$, a possible approach is to lift *locally* the plane CWT from the tangent plane along the local normal, thus defining a genuine notion of local dilation on $\mathcal{M}$. In this way, one obtains a local CWT on $\mathcal{M}$. It remains to glue the local patches by standard methods of differential geometry, but that is not always easy [12].

5. *A unified point of view*

   Let us come back for a moment to the group structure underlying the CWT. In 1-D, the group is the affine group of the line $G_{\mathrm{aff}}^{\mathbb{R}} = \mathbb{R} \rtimes \mathbb{R}_*$. In 2-D, one gets the similitude group of the plane $\mathrm{SIM}(2) = \mathbb{R}^2 \rtimes (\mathbb{R}_*^+ \times \mathrm{SO}(2))$. Now, using polar coordinates, the right hand factor $\mathbb{R}_*^+ \times \mathrm{SO}(2)$ may be identified with the pointed plane $\mathbb{R}_*^2$. Further identifying the real plane $\mathbb{R}^2$ with the complex plane $\mathbb{C}$, we finally obtain that the group underlying the 2-D CWT is $G_{\mathrm{aff}}^{\mathbb{C}} = \mathbb{C} \rtimes \mathbb{C}_*$.

This approach may be pursued one step further, following Ali [6]. The idea is to define *quaternionic wavelets*. Let $\mathbb{H}$ denote the field of all quaternions and $\mathbb{H}_*$ the group (under quaternionic multiplication) of all invertible quaternions. It turns out that the group $\mathbb{H}_*$ is isomorphic to the *affine SU(2) group*, i.e., $\mathbb{R}^+_* \times \mathrm{SU}(2)$, that is, the group SU(2) together with all (non-zero) dilations. As a set $\mathbb{H}_* \simeq \mathbb{R}^+_* \times \mathbb{S}^3$, where $\mathbb{S}^3$ is the surface of the unit sphere in $\mathbb{R}^4$, or more simply, $\mathbb{H}_* \simeq \mathbb{R}^4 \setminus \{\mathbf{0}\}$. Exactly as in the two other cases, $\mathbb{H}_*$ acts transitively on $\mathbb{H}$ and from this one can construct quaternionic wavelets, with underlying group $G^{\mathbb{H}}_{\mathrm{aff}} = \mathbb{H} \rtimes \mathbb{H}_*$.

In this way, one obtains a useful and elegant approach to the three cases, with the group $G^{\mathbb{K}}_{\mathrm{aff}} = \mathbb{K} \rtimes \mathbb{K}_*$, where $\mathbb{K} = \mathbb{R}, \mathbb{C}$ or $\mathbb{H}$. We refer to [6] for the development of this type of wavelets.

11. Spatio-temporal wavelets and motion estimation

There exist many methods for motion estimation : Optical flow, Block matching, Phase difference, ... An alternative approach is provided by the motion-tuned continuous wavelet transform, developed by M. Duval-Destin (1991), R.Murenzi (1992) and F. Mujica (1999). The idea is to adapt to the (2D+T) space-time the general formalism of the continuous wavelet transform on a manifold, described in Section 10.1 above. The rest of this section is based on our paper [15].

Technically speaking, one builds a time dependent wavelet, separable in frequency space:

$$\widehat{\psi}_{ST}(\vec{k}, \omega) = \underbrace{\widehat{\psi}_S(k_x, k_y)}_{\text{2-D wavelet}} \cdot \underbrace{\widehat{\psi}_T(\omega)}_{\text{1-D wavelet}} \,. \tag{11.1}$$

Then one acts on it by the space-time (2D+T) group, containing space and time translations, space and time dilations, space rotations, namely, $G = \mathrm{SIM}(2) \times G^+_{\mathrm{aff}}$, via the usual unitary irreducible, square integrable, representations in the space of signals (image sequences) $L^2(\mathbb{R}^2 \times \mathbb{R},\, \mathrm{d}\vec{x}\ \mathrm{d}t)$. Writing the Fourier transform as

$$\widehat{s}(\vec{k}, \omega) = (2\pi)^{-3/2} \iint_{\mathbb{R}^2 \times \mathbb{R}} e^{-i(\vec{k}\cdot\vec{x} + \omega t)} s(\vec{x}, t)\, \mathrm{d}\vec{x}\ \mathrm{d}t\,,$$

the Fourier space is simply $L^2(\mathbb{R}^2 \times \mathbb{R},\, \mathrm{d}\vec{k}\ \mathrm{d}\omega)$.

Finally, one replaces the separate space and time dilations a_s, a_t by a *global* dilation a and a *speed tuning* parameter c.

The principle underlying motion estimation is that speed detection and quantization is done in the Fourier space, because the wavelet measures the inclination of the signal spectrum. Indeed the latter increases with speed, as shown in Fig. 5: a static object $\widehat{s}$ lives in the plane $\omega = 0$ of zero frequency, whereas an object $\widehat{s}$ moving with constant speed $\vec{v}$ lives in the plane $\vec{k} \cdot \vec{v} + \omega = 0$.

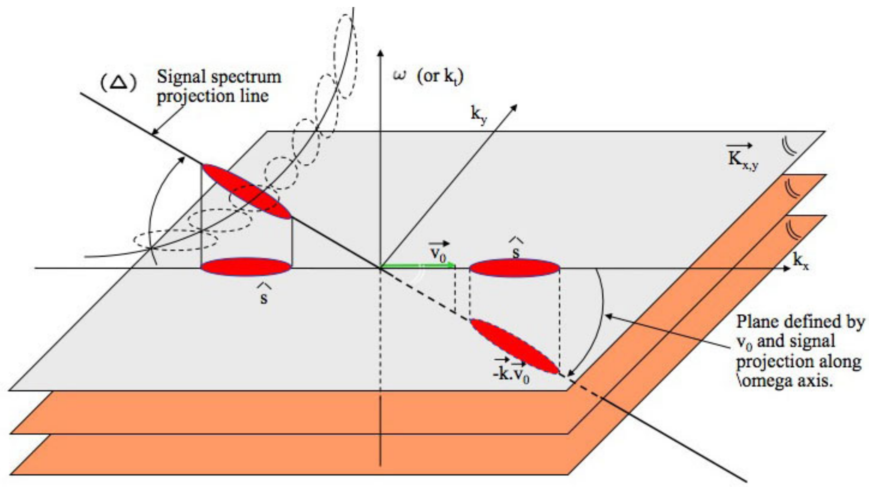


FIGURE 5. Inclination of a signal spectrum as a function of speed

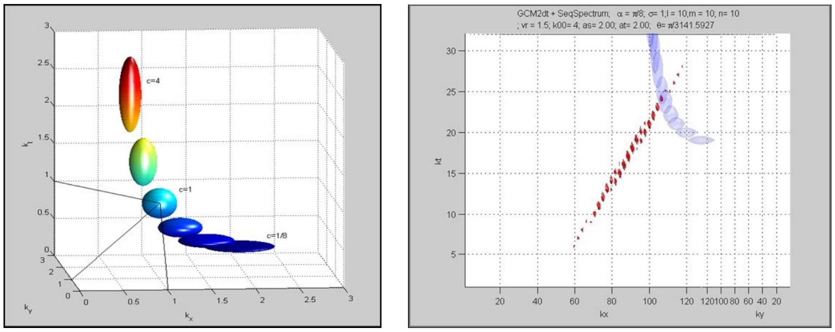
The next step is to identify the unitary motion operators acting in the Fourier space $L^2(\mathbb{R}^2 \times \mathbb{R},\, \mathrm{d}\vec{k}\, \mathrm{d}\omega)$, namely,

$$\begin{array}{lrcl}
\text{Dilation :} & [\widehat{D}^a\, \widehat{\psi}](\vec{k},\omega) & = & a^{3/2}\widehat{\psi}(a\vec{k}, a\,\omega) \\
\text{Translation :} & [\widehat{T}^{\vec{b},\tau}\, \widehat{\psi}](\vec{k},\omega) & = & e^{-i(\vec{k}\cdot\vec{b}+\omega\tau)}\widehat{\psi}(\vec{k},\omega) \\
\text{Rotation :} & [\widehat{R}^{\theta}\, \widehat{\psi}](\vec{k},\omega) & = & \widehat{\psi}(r_{-\theta}\vec{k},\omega) \\
\text{Speed tuning :} & [\widehat{\Lambda}^{c}\, \widehat{\psi}](\vec{k},\omega) & = & \widehat{\psi}(c^{q}\vec{k}, c^{-p}\omega)
\end{array}$$

In the definition of speed tuning, we have the additional constraints that $\widehat{\Lambda}^c$ must map the $\vec{v}_o$-plane, $\vec{k}\cdot\vec{v}_o + \omega = 0$, into the $c\vec{v}_o$-plane and must be unitary, which implies that $p = 2/3$ and $q = 1/3$. This correspond to the psycho-visual effect. If an object moves fast, only *large* details can be detected, but, if it moves slowly, then *small* details can be detected. Hence, a high speed object must be large to be "captured" and a low speed object can be small. Therefore wavelets must be speed-tuned (distorted and elongated) to capture a moving objet, which means that wavelets move on a hyperbola-like curve with increasing speed, as seen on Fig. 6 (a). We show on Fig. 6 (b) the speed analysis of an object moving at constant speed. Capture is achieved when the signal spectrum (red) intersects the family of speed-tuned wavelets (blue).

The next step is to choose adequate wavelets for space and time components. The standard 2D+T wavelet is the Duval-Destin-Murenzi (DDM) wavelet

$$\widehat{\psi}_{\rm DDM}(\vec{k},\omega) = \underbrace{\widehat{\psi}_{\rm M}(k_x,k_y)}_{\text{2-D Morlet wavelet}} \cdot \underbrace{\widehat{\psi}_{\rm M}(\omega)}_{\text{1-D Morlet}}$$



(a) (b)

Figure 6. (a) The hyperbola of speed-tuned wavelets. (b) Speed analysis of an object moving at constant speed.

where the 1-D Morlet wavelet is

$$\widehat{\psi}_{\mathrm{M}}(\omega) = \exp(-\tfrac{1}{2}(\omega-\omega_0)^2) - \widehat{h}(\omega),$$

with the correction term $\widehat{h}(\omega)$ negligible in practice (for $\omega_0 \gtrsim 5.5$).

Since the 2-D Morlet wavelet has poor selectivity properties, we replace it by a *Gaussian-conical wavelet*, as given in (9.13), and get a GCM 2D+T wavelet

$$\widehat{\psi}^{\mathrm{GCM}}_{lm}(\vec{k},\omega) = \begin{cases} \underbrace{\widehat{\psi}^{\mathrm{GC}}_{lm}(k_x,k_y)}_{\text{2-D Gaussian-Conical}} \cdot \underbrace{\widehat{\psi}_{\mathrm{M}}(\omega)}_{\text{1-D Morlet}}, & \vec{k} \in \mathcal{C}(-\alpha,\alpha), \\ 0, \ \text{otherwise}, \end{cases}$$

Explicitly

$$\widehat{\psi}^{\mathrm{GCM}}_{lm}(\vec{k},\omega) = \begin{cases} (\vec{k}\cdot\vec{e}_{-\widetilde{\alpha}})^l\,(\vec{k}\cdot\vec{e}_{\widetilde{\alpha}})^m e^{-\frac{\sigma}{2}(k_x-\chi(\sigma))^2}\, e^{-\frac{1}{2}(\omega-\omega_0)^2}, \ \vec{k} \in \mathcal{C}(-\alpha,\alpha), \\ 0, \ \text{otherwise} \end{cases}$$

and the so-called center correction term $\chi(\sigma) = \sqrt{l+m}\frac{\sigma-1}{\sigma}$ controls the radial support of the wavelet $\widehat{\psi}^{\mathrm{GCM}}_{lm}$.

Several experiments [15] prove that the GCM 2D+T wavelet has a better angular resolving power than the DDM 2D+T wavelet, it shows an extreme efficiency in directional speed selectivity down to very small angle apertures, it has a good capacity of radial stability and adjustment with respect to variation of the aperture and, last but not least, it has a good stability with respect to noise.

In conclusion, we may say that the GCM 2D+T wavelet is a highly directionally selective speed-tuned wavelet, which provides a much more powerful tool than the DDM 2D+T wavelet for the recognition and tracking of spectral signatures.

References

[1] S.T. Ali and M. Engliš, *Quantization methods: a guide for physicists and analysts*, Rev. Math. Phys. **17** (2005), 391–490.

[2] S.T. Ali, J-P. Antoine, J-P. Gazeau and U.A. Mueller, *Coherent states and their generalizations: A mathematical overview*, Reviews Math. Phys. **7** (1995), 1013–1104.

[3] S.T. Ali, J-P. Antoine and J-P. Gazeau, *Coherent States, Wavelets and Their Generalizations*, 2nd ed., Springer-Verlag, New York et al., 2014.

[4] S.T. Ali, T. Bhattacharyya and S.S. Roy, *Coherent states on Hilbert modules*, J. Phys. A: Math. Theor. **44** (2011), 275202.

[5] S.T. Ali, F. Bagarello and J-P. Gazeau, *Quantizations from reproducing kernel spaces*, Ann. Phys. **332** (2013), 127–142.

[6] S.T. Ali and K. Thirulogasanthar, *The quaternionic affine group and related continuous wavelet transforms on complex and quaternionic Hilbert spaces*, preprint (2014); arXiv:1402.3109v1.

[7] J-P. Antoine and P. Vandergheynst, *Wavelets on the n-sphere and related manifolds,* J. Math. Phys. **39** (1998), 3987–4008.

[8] J-P. Antoine and P. Vandergheynst, *Wavelets on the 2-sphere: A group-theoretical approach,* Appl. Comput. Harmon. Anal. **7** (1999), 262–291.

[9] J-P. Antoine, L. Demanet, L. Jacques, and P. Vandergheynst, *Wavelets on the sphere: Implementation and approximations,* Appl. Comput. Harmon. Anal. **13** (2002), 177–200.

[10] J-P. Antoine, R. Murenzi, P. Vandergheynst and S.T. Ali, *Two-dimensional Wavelets and Their Relatives*, Cambridge University Press, Cambridge (UK), 2004.

[11] J-P. Antoine, I. Bogdanova, and P. Vandergheynst, *The continuous wavelet transform on conic sections,* Int. J. Wavelets, Multires. and Inform. Proc., **6** (2007) 137–156.

[12] J-P. Antoine, D. Roşca, and P. Vandergheynst, *Wavelet transform on manifolds: Old and new approaches,* Appl. Comput. Harmon. Anal. **28** (2010), 189–202.

[13] I. Bogdanova, P. Vandergheynst, J.-P. Antoine, L. Jacques, and M. Morvidone, *Stereographic wavelet frames on the sphere,* Appl. Comput. Harmon. Anal. **19** (2005), 223–252.

[14] I. Bogdanova, P. Vandergheynst, and J-P. Gazeau, *Continuous wavelet transform on the hyperboloid,* Appl. Comput. Harmon. Anal. **23** (2007), 286–306.

[15] P. Brault and J-P. Antoine, *A spatio-temporal Gaussian-Conical wavelet with high aperture selectivity for motion and speed analysis*, Appl. Comput. Harmon. Anal. **34** (2013), 148–161.

[16] M. Calixto, J. Guerrero, and D. Roşca, *Wavelet transform on the torus: a group-theoretical approach*, Appl. Comput. Harmon. Anal. **37** (2014), in press; http://dx.doi.org/10.1016/j.acha.2014.03.001

[17] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, 1992.

[18] D. Gabor, *Theory of communication,* J. Inst. Electr. Engrg. (London) **93** (1946), 429–457.

[19] J-P. Gazeau, *Coherent States in Quantum Physics*, Wiley-VCH, Berlin, 2009.

[20] R. Gilmore, *Geometry of symmetrized states*, Ann. Phys. (NY) **74** (1972), 391–463; *On properties of coherent states*, Rev. Mex. Fis. **23** (1974), 143–187.

[21] A. Grossmann, J. Morlet and T. Paul, *Integral transforms associated to square integrable representations. I. General results,* J. Math. Phys. **26** (1985), 2473–2479; *II. Examples,* Ann. Inst. H. Poincaré **45** (1986), 293–309.

[22] C. Heil and D. Walnut (eds.), *Fundamental Papers in Wavelet Theory*, Princeton University Press, Princeton, NJ, 2006.

[23] E.C. Lance, *Hilbert C*-modules*, London Math. Soc. Lect. Notes Series 210, Cambridge UP, Cambridge, 1995.

[24] S.G. Mallat, *A Wavelet Tour of Signal Processing, 3nd ed. : The Sparse Way*, Academic Press, San Diego, 2008.

[25] D. Marr, *Vision*, Freeman, San Francisco, 1982.

[26] A.M. Perelomov, *Coherent states for arbitrary Lie group*, Commun. Math. Phys. **26** (1972), 222–236.

[27] D. Roşca and J-P. Antoine, *Locally supported orthogonal wavelet bases on the sphere via stereographic projection,* Math. Probl. Engineering, Vol. **2009** (2009), 124904.

[28] K. Thirulogasanthar and S.T. Ali, *A class of vector coherent states defined over matrix domains*, J. Math. Phys. **44** (2003), 5070–5083.

[29] K. Thirulogasanthar, G. Honnouvo and A. Krzyżak, *Coherent states and Hermite polynomials on quaternionic Hilbert spaces*, J. Phys. A: Math. Theor. **43** (2010), 385205.

[30] J. Ville, *Théorie et applications de la notion de signal analytique,* Câbles et Transm. **2ème A** (1948), 61–74 .

[31] E.P. Wigner, *On the quantum correction for thermodynamic equilibrium,* Phys. Rev. **40** (1932), 749–759 .

Jean-Pierre Antoine
Institut de Recherche en Mathématique et Physique
Université catholique de Louvain
B - 1348 Louvain-la-Neuve
Belgium
e-mail: jean-pierre.antoine@uclouvain.be