Chapter 7

Equivariant Convolutional Neural Networks

7.1 Steerable Families

Since it is not practical to use the definition of cross-correlation involving integration over the group $G = \mathbb{R}^d \rtimes H$ we go back to the notion of lifted correlation.

It is more convenient to assume that the semi-direct product $G = \mathbb{R}^d \rtimes H$ is defined by an action of H on $\mathbb{R}^d$ by automorphisms so that elements of G are denoted as pairs $(x, h) \in \mathbb{R}^d \times H$.

Then for any function $f \in \mathrm{L}^2(\mathbb{R}^d)$ and any correlation kernel $k \in \mathrm{L}^1(\mathbb{R}^d)$ with compact support, the *lifted correlation* $k \,\widetilde{\star}\, f$ is defined by $(*_{10'})$, namely

$$(k \,\widetilde{\star}\, f)(x,h) = \int_{\mathbb{R}^d} f(t)k(h^{-1}\cdot(t-x))\, dt, \, (x,h) \in \mathbb{R}^d \times H.$$

Observe that $k \,\widetilde{\star}\, f$ is a function with domain $\mathbb{R}^d \times H$.

Computing $(k \,\widetilde{\star}\, f)(x,h)$ requires sampling the group H, which is too expensive if $d \geq 3$.

A way around this problem is to express the kernels k in terms of a basis of "steerable functions." Intuitively this means using some kind of generalized harmonic functions.

In our case we need to find bases of functions in $\mathrm{L}^2(\mathbb{R}^d)$ that are H-steerable.

In applications H is a compact group so as we will see shortly, we can use the Peter–Weyl theorem, actually Version II, namely Theorem 4.4, to find steerable bases.

The problem is the presence of the term $k(h^{-1} \cdot (t-x))$ in the integral defining $k \widetilde{\star} f$.

The key point is that if we can express the kernel k as a linear combination of linearly independent functions $Y_1, \ldots, Y_L$ in $\mathrm{L}^2(\mathbb{R}^d)$ that are "nice," which means that for every $h \in H$ and every $x \in \mathbb{R}^d$, each $Y_j(h^{-1} \cdot x)$ can be expressed as a linear combination of $Y_1(x), \ldots, Y_L(x)$, then it is possible to express $(k \widetilde{\star} f)(x, h)$ in a linear fashion in terms of the vector

$$f^Y(x) = \int_{\mathbb{R}^d} f(t) Y(t-x)\, dt,$$

where $Y(x)$ denotes the column vector

$$Y(x) = \begin{pmatrix} Y_1(x) \\ \vdots \\ Y_L(x) \end{pmatrix} \in \mathbb{C}^L.$$

So let us assume that for every $h \in H$, there is an invertible matrix $\Sigma(h) \in \mathbf{GL}(L, \mathbb{C})$, such that

$$\begin{pmatrix} Y_1(h^{-1} \cdot x) \\ Y_2(h^{-1} \cdot x) \\ \vdots \\ Y_L(h^{-1} \cdot x) \end{pmatrix} = \begin{pmatrix} \Sigma(h)_{11} & \Sigma(h)_{2,1} & \cdots & \Sigma(h)_{L1} \\ \Sigma(h)_{12} & \Sigma(h)_{2,2} & \cdots & \Sigma(h)_{L2} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma(h)_{1L} & \Sigma(h)_{2,L} & \cdots & \Sigma(h)_{LL} \end{pmatrix} \begin{pmatrix} Y_1(x) \\ Y_2(x) \\ \vdots \\ Y_L(x) \end{pmatrix},$$

or more concisely,

$$Y(h^{-1} \cdot x) = \Sigma(h)^{\top} Y(x), \quad x \in \mathbb{R}^d. \qquad \text{(steer1)}$$

If Equation (steer1) holds we say that $(Y_1, \ldots, Y_L)$ is an *H-steerable family* (or *H-steerable basis*). For short, we often drop H.

In fact, we will see later that the map $\Sigma \colon H \to \mathbf{U}(L)$ is a *representation of H*.

The reason for using $\Sigma(h)^\top$ instead of $\Sigma(h)$ is technical and will become clear later when we explain how to create steerable families.

The notion of steerability occured first in the seminal paper of Freeman and Adelson [20].

Next assume that the kernel k can be expressed as a linear combination of the Y_i using some coefficients $w_i \in \mathbb{C}$ that we call *weights*.

Let us write

$$k(x; w) = \sum_{i=1}^{L} \overline{w_i}\, Y_i(x) = w^* Y(x),\ x \in X, \qquad \text{(kw1)}$$

where $w \in \mathbb{C}^L$ is the column vector consisting of the w_i.

The reason for using conjugate weights will become apparent in the computation below.

Let us compute $k(h^{-1} \cdot x; w)$. Since Y is a steerable family, we have

$$\begin{aligned} k(h^{-1} \cdot x; w) &= w^* Y(h^{-1} \cdot x) = w^* \Sigma(h)^\top Y(x) \\ &= (\overline{\Sigma(h)} w)^* Y(x) = k(x; \overline{\Sigma(h)} w), \end{aligned}$$

namely

$$k(h^{-1} \cdot x; w) = k(x; \overline{\Sigma(h)} w). \qquad \text{(kw2)}$$

So the new kernel is obtained by simply modifying the weights using the matrix $\overline{\Sigma(h)}$.

And now let us compute the *lifted correlation* $k \mathbin{\widetilde{\star}} f$ given by

$$(k \mathbin{\widetilde{\star}} f)(x, h) = \int_{\mathbb{R}^d} f(t) k(h^{-1} \cdot (t - x); w)\, dt,$$

with $(x, h) \in \mathbb{R}^d \rtimes H$.

Using the fact that k is a steerable family we have

$$\begin{aligned}(k \mathbin{\widetilde{\star}} f)(x, h) &= \int_{\mathbb{R}^d} f(t) k(h^{-1} \cdot (t - x); w)\, dt \\ &= \int_{\mathbb{R}^d} f(t) w^* \Sigma(h)^\top Y(t - x)\, dt \\ &= w^* \Sigma(h)^\top \int_{\mathbb{R}^d} f(t) Y(t - x)\, dt.\end{aligned}$$

Let f^Y be the function $f^Y \colon \mathbb{R}^d \to \mathbb{C}^L$ given by

$$f^Y(x) = \begin{pmatrix} \int_{\mathbb{R}^d} f(t)Y_1(t-x)\,dt \\ \vdots \\ \int_{\mathbb{R}^d} f(t)Y_L(t-x)\,dt \end{pmatrix} = \int_{\mathbb{R}^d} f(t)Y(t-x)\,dt. \tag{f^Y}$$

Then using the trick that for any two column vectors $u, v \in \mathbb{C}^n$, we have

$$u^\top v = \mathrm{tr}(vu^\top),$$

we obtain

$$\begin{aligned} (k \,\widetilde{\star}\, f)(x,h) &= \int_{\mathbb{R}^d} f(t)k(h^{-1}\cdot(t-x); w)\,dt \\ &= w^*\Sigma(h)^\top \int_{\mathbb{R}^d} f(t)Y(t-x)\,dt \\ &= w^*\Sigma(h)^\top f^Y(x) = \mathrm{tr}(f^Y(x)w^*\,\Sigma(h)^\top), \end{aligned}$$

where we use the identity

$$u^\top v = \mathrm{tr}(vu^\top)$$

with $u = (w^*\Sigma(h)^\top)^\top$ and $v = f^Y(x)$.

Observe that

$$\widehat{f}(x) = f^Y(x)w^*$$

is an $L \times L$ matrix that can be thought of as *some kind of Fourier coefficients of f*.

The formula

$$(k \,\widetilde{\star}\, f)(x, h) = \mathrm{tr}\left(f^Y(x)w^*\, \Sigma(h)^\top\right) = \mathrm{tr}\left(\widehat{f}(x)\, \Sigma(h)^\top\right),$$

shows that $(k \,\widetilde{\star}\, f)(x, h)$ is *similar to a Fourier cotransform with respect to the representation* $\Sigma^\top$, where $\widehat{f}(x)/L$ plays the role of $\mathcal{F}(f)(\rho)$ and $\Sigma^\top$ plays the role of M_ρ (see Formula (FI) in Section 4.7 and Theorem 4.26), except that $\Sigma^\top$ is not necessarily irreducible.

The function $\widehat{f}\colon \mathbb{R}^d \to \mathrm{M}_L(\mathbb{C})$ given by $\widehat{f}(x) = f^Y(x)w^*$ is a *matrix-valued function*.

What we have gained is that when we compute the integral

$$f^Y(x) = \int_{\mathbb{R}^d} f(t)Y(t-x)\,dt,$$

we incorporate all the information about the action of H on $\mathbb{R}^d$ into f^Y *without having to sample* the group H.

The functions $(Y_1, \ldots, Y_L)$ package all the information about the group H needed to compute the essential part of $(k \mathbin{\widetilde{\star}} f)(x, h)$.

The outer product $\widehat{f}(x) = f^Y(x)w^*$ incorporates all the information in the kernel k using the weight vector w.

Computing

$$(k \,\widetilde{\star}\, f)(x, h) = \mathrm{tr}\left(\widehat{f}(x)\,\Sigma(h)^\top\right) = \mathrm{tr}\left(\Sigma(h)\widehat{f}(x)^\top\right)$$

is then very cheap, since it is a linear operation only involving the matrix $\Sigma(h)$.

Another important observation is that starting with an input function $f \in \mathrm{L}^2(\mathbb{R}^d)$, the lifted correlation $k \,\widetilde{\star}\, f$ is a *scalar-valued* function (with codomain $\mathbb{C}$) defined on the *augmented domain* $\mathbb{R}^d \times H$, but $\widehat{f}$ is a *vector-valued* function from $\mathbb{R}^d$ to the *augmented codomain* $\mathrm{M}_L(\mathbb{C})$.

The group $\mathbb{R}^d \times H$ acts on the domain $\mathbb{R}^d$ of $\widehat{f}$, and the group H acts on its codomain $\mathrm{M}_L(\mathbb{C})$ in terms of the representation Σ by multiplication on the left by $\Sigma(h)$.

This is one of the motivations for introducing certain vector-valued functions called *feature fields*, discussed in the Section 7.3.

The notion of steerability is easily generalized to any measure space X such that $\mathrm{L}^2(X)$ is separable and H acts continuously on X.

For example, any locally compact, metrizable, separable space X equipped with a σ-regular, locally finite, Borel measure μ will do; see Vol I, Theorem @@@7.11.

Definition 7.1. Let X be any measure space such that $\mathrm{L}^2(X)$ is separable and H acts continuously on X. Some linerarly independent functions $(Y_1, \ldots, Y_L)$ in $\mathrm{L}^2(X)$ form an *H-steerable family* (or *H-steerable basis*) if there is a representation $\Sigma\colon H \to \mathbf{U}(L)$ such that

$$Y(h^{-1} \cdot x) = \Sigma(h)^\top Y(x), \quad h \in H,\, x \in X, \quad \text{(steer2)}$$

where $Y(x)$ denotes the column vector

$$Y(x) = \begin{pmatrix} Y_1(x) \\ \vdots \\ Y_L(x) \end{pmatrix} \in \mathbb{C}^L.$$

This more general notion will be needed in Section **??** to construct equivariant kernels.

The simplest example (simpler that $X = \mathbb{R}^2$) is the circle, $X = S^1$, with $H = \mathbf{SO}(2)$, the group of rotations in the plane.

Example 7.1. Let $H = \mathbf{SO}(2)$ and $X = S^1 \approx \mathbf{SO}(2)$. For any L-tuple of integers $(n_1, \ldots, n_L)$, we claim that

$$Y(\alpha) = (e^{-in_1\alpha}, \ldots, e^{-in_L\alpha})$$

is a steerable family (where the expression on the right-hand side denotes a column vector).

As we saw in Proposition 2.12, every unitary representation $\Sigma\colon \mathbf{SO}(2) \to \mathbf{U}(L)$ is given by a matrix of the form

$$\Sigma(\alpha) = \begin{pmatrix} e^{ik_1\alpha} & 0 & \dots & 0 \\ 0 & e^{ik_2\alpha} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & e^{ik_L\alpha} \end{pmatrix}$$

with $k_1, \dots, k_L \in \mathbb{Z}$, so if we pick $k_j = n_j$, for $j = 1, \dots, L$ and

$$Y_j(\alpha) = e^{-in_j\alpha},$$

since

$$Y_j(\alpha - \theta) = e^{-in_j(\alpha-\theta)} = e^{in_j\theta}e^{-in_j\alpha} = e^{in_j\theta}Y_j(\alpha),$$

we see that

$$\begin{pmatrix} Y_1(\alpha - \theta) \\ \vdots \\ Y_L(\alpha - \theta) \end{pmatrix} = \begin{pmatrix} e^{in_1\theta} & 0 & \dots & 0 \\ 0 & e^{in_2\theta} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & e^{in_L\theta} \end{pmatrix} \begin{pmatrix} Y_1(\alpha) \\ \vdots \\ Y_L(\alpha) \end{pmatrix},$$

which confirms that $(Y_1(\alpha) = e^{-in_1\alpha}, \dots, Y_L(\alpha) = e^{-in_L\alpha})$ is a steerable family (again, the expression on the right-hand side denotes a column vector).

7.2 Construction of H-Steerable Families

We now present a method for finding steerable families on a space X as above equipped with a continuous action of a compact group H.

The trick is to consider the unitary representation $V\colon H \to \mathbf{U}(\mathrm{L}^2(X))$ given by

$$(V(h)f)(x) = f(h^{-1} \cdot x), \quad h \in H,\ f \in \mathrm{L}^2(X),\ x \in X. \tag{V}$$

According to the Peter–Weyl theorem, Version II, the space $\mathrm{L}^2(X)$ is the Hilbert sum of closed subspaces E_ρ with $\rho \in R(H)$ (which may be reduced to zero), where E_ρ is the projection of $\mathrm{L}^2(X)$ under the projection π_ρ^V given by

$$\pi_\rho^V(f) = n_\rho \int_H \overline{\chi_\rho(h)}(V(h)f)\, d\lambda(h),$$

where $f \in \mathrm{L}^2(X)$ and λ is a left Haar measure on H.

First we need to take care of a technicality.

As stated the theorem involves the projection $\pi_\rho^V(f)$ of a function $f \in \mathrm{L}^2(X)$ and it is defined as a weak integral.

For our purposes we need a formula definining $(\pi_\rho^V(f))(x)$ for every $x \in H$.

This can be achieved as follows.

By definition of $\pi_\rho^V(f)$ as a weak integral it is the unique function (given by the Riesz representation theorem, Theorem 4.7(2)) such that

$$\langle \pi_\rho^V(f), g\rangle = n_\rho \int_H \overline{\chi_\rho(h)} \langle V(h)f, g\rangle \, d\lambda(h)$$

for all $g \in \mathrm{L}^2(X)$, and using the definition of the inner product on $\mathrm{L}^2(X)$ and Fubini the above is expressed as

$$\begin{aligned}
&\int_X (\pi_\rho^V(f))(x)\overline{g(x)}\, d\mu_X(x) \\
&= n_\rho \int_H \overline{\chi_\rho(h)} \int_X (V(h)f)(x)\overline{g(x)}\, d\mu_X(x)\, d\lambda(h) \\
&= n_\rho \int_X \left(\int_H \overline{\chi_\rho(h)} (V(h)f)(x)\, d\lambda(h) \right) \overline{g(x)}\, d\mu_X(x),
\end{aligned}$$

and since it holds for all $g \in \mathrm{L}^2(X)$, we must have

$$
\begin{aligned}
(\pi_\rho^V(f))(x) &= n_\rho \int_H \overline{\chi_\rho(h)}(V(h)f)(x)\, d\lambda(h) \\
&= n_\rho \int_H \overline{\chi_\rho(h)} f(h^{-1}\cdot x)\, d\lambda(h). \qquad (\pi_\rho^V)
\end{aligned}
$$

So the projection $\pi_\rho^V(f)$ of the function $f \in \mathrm{L}^2(X)$ *can be defined pointwise* by (π_ρ^V), but it is not obvious *a priori* that this yields a function in $\mathrm{L}^2(X)$, which is guaranteed by the weak integral argument.

Going back to Peter–Weyl II, each subspace E_ρ is a finite or countably infinite Hilbert sum of d_ρ (where $d_\rho = \infty$ is possible) closed finite-dimensional subspaces $E_\rho^{k_\rho}$ ($1 \leq k_\rho \leq d_\rho$) such that for every ρ and every k_ρ,

each subrepresentation $V_\rho^{k_\rho}\colon H \to \mathbf{U}(E_\rho^{k_\rho})$ is equivalent to the irreducible representation $M_\rho\colon H \to \mathbf{U}(\mathbb{C}^{n_\rho})$.

Thus there are linear isomorphisms $\theta_\rho^{k_\rho}\colon E_\rho^{k_\rho} \to \mathbb{C}^{n_\rho}$ such that the following diagrams commute

$$\begin{array}{ccc} E_\rho^{k_\rho} & \xrightarrow{\theta_\rho^{k_\rho}} & \mathbb{C}^{n_\rho} \\ {\scriptstyle V_\rho^{k_\rho}(h)}\downarrow & & \downarrow{\scriptstyle M_\rho(h)} \\ E_\rho^{k_\rho} & \xrightarrow[\theta_\rho^{k_\rho}]{} & \mathbb{C}^{n_\rho} \end{array}$$

for all $h \in H$.

Since $(V(h)f)(x) = f(h^{-1} \cdot x)$, we have

$$f(h^{-1} \cdot x) = (V_\rho^{k_\rho}(h)f)(x), \quad h \in H,\ f \in E_\rho^{k_\rho},\ x \in X, \tag{steer3}$$

for all $\rho \in R(H)$ and all k_ρ.

If we pick an orthonormal basis (orthogonal works too) $(Y^1_{\rho,k_\rho}, \ldots, Y^{n_\rho}_{\rho,k_\rho})$ in each $E^{k_\rho}_\rho$ so that the family $(Y^j_{\rho,k_\rho})_{\rho\in R(H), 1\leq k_\rho\leq d_\rho, 1\leq j\leq n_\rho}$ is a Hilbert basis of $\mathrm{L}^2(X)$, then there is an $n_\rho \times n_\rho$ unitary matrix $M^{\rho,k_\rho}(h)$ representing the linear map $V^{k_\rho}_\rho(h)$ with respect to the basis $(Y^1_{\rho,k_\rho}, \ldots, Y^{n_\rho}_{\rho,k_\rho})$ defined by

$$Y^j_{\rho,k_\rho}(h^{-1}\cdot x) = \sum_{i=1}^{n_\rho} M^{\rho,k_\rho}_{ij}(h) Y^i_{\rho,k_\rho}(x).$$

If we stack the $Y^i_{\rho,k_\rho}(x)$ into a column vector $Y_{\rho,k_\rho}(x)$ and the $Y^j_{\rho,k_\rho}(h^{-1}\cdot x)$ into a column vector $Y_{\rho,k_\rho}(h^{-1}\cdot x)$, we can write

$$Y_{\rho,k_\rho}(h^{-1}\cdot x) = (M^{\rho,k_\rho}(h))^\top Y_{\rho,k_\rho}(x). \tag{steer4}$$

Remark: The presence of the transposition is the familiar artifact of linear algebra caused by the fact that $Y_{\rho,k_\rho}(x)$ is a column vector.

Replacing h by h^{-1} in (steer4) we get

$$\begin{aligned} Y_{\rho,k_\rho}(h \cdot x) &= (M^{\rho,k_\rho}(h^{-1}))^\top Y_{\rho,k_\rho}(x) \\ &= ((M^{\rho,k_\rho}(h))^*)^\top Y_{\rho,k_\rho}(x) \\ &= \overline{M^{\rho,k_\rho}(h)} Y_{\rho,k_\rho}(x), \end{aligned}$$

so conjugating on both sides we get

$$\overline{Y_{\rho,k_\rho}}(h \cdot x) = M^{\rho,k_\rho}(h) \overline{Y_{\rho,k_\rho}}(x). \tag{steer5}$$

Observe that the unitary representation $V_\rho^{k_\rho} \colon H \to \mathbf{U}(E_\rho^{k_\rho})$ define a representation in matrix form $M^{\rho,k_\rho} \colon H \to \mathbf{U}(\mathbb{C}^{n_\rho})$ equivalent to the irreducible representation M_ρ. The equations (steer4) express the fact that the basis functions $(Y^1_{\rho,k_\rho}, \ldots, Y^{n_\rho}_{\rho,k_\rho})$ are steerable.

Definition 7.2. If H is a compact group and X is a locally compact, metrizable, separable space equipped with a σ-regular, locally finite, Borel measure μ, given any continuous action of H on X, some linerarly independent functions $(Y_1, \ldots, Y_L)$ in $\mathrm{L}^2(X)$ form an *H-steerable family* (or *H-steerable basis*) if there is a representation $\Sigma\colon H \to \mathbf{U}(L)$ such that

$$Y(h^{-1} \cdot x) = \Sigma(h)^{\top} Y(x), \quad h \in H,\ x \in X, \qquad \text{(steer6)}$$

or equivalently

$$\overline{Y}(h \cdot x) = \Sigma(h)\, \overline{Y}(x), \quad h \in H,\ x \in X, \qquad \text{(steer7)}$$

where $Y(x)$ denotes the column vector

$$Y(x) = \begin{pmatrix} Y_1(x) \\ \vdots \\ Y_L(x) \end{pmatrix} \in \mathbb{C}^L.$$

Remark: Steerability as defined above is equivalent to the notion of steerability as defined in Lang and Weiler [31].

In Cesa, Lang and Weiler [7] as well as Bekkers [1], the notion of steerability is defined using $Y(h \cdot x)$ instead of $Y(h^{-1} \cdot x)$.

We pass from one version to the other by conjugation of the functions. In the papers mentioned above, steerable families are also called *harmonic basis functions*.

Due to its importance, the preceding discussion is summarized in the following theorem.

Theorem 7.1. *Let X be a locally compact, metrizable, separable space equipped with a σ-regular, locally finite, Borel measure μ. If H is a compact group acting continuously on X (not necessarily in a transitive fashion), consider the unitary representation $V\colon H \to \mathbf{U}(\mathrm{L}^2(X))$ given by*

$$(V(h)f)(x) = f(h^{-1} \cdot x), \quad h \in H,\ f \in \mathrm{L}^2(X),\ x \in X.$$

According to the Peter–Weyl theorem, Version II, the space $\mathrm{L}^2(X)$ is the Hilbert sum of closed subspaces E_ρ with $\rho \in R(H)$ (which may be reduced to zero), where E_ρ is the projection of $\mathrm{L}^2(X)$ under the projection π_ρ^V given by

$$(\pi_\rho^V(f))(x) = n_\rho \int_H \overline{\chi_\rho(h)} f(h^{-1} \cdot x)\, d\lambda(h),$$

where $f \in \mathrm{L}^2(X)$, $x \in X$, and λ is a left Haar measure on H.

Each subspace E_ρ is a finite or countably infinite Hilbert sum of d_ρ (where $d_\rho = \infty$ is possible) closed finite-dimensional subspaces $E_\rho^{k_\rho}$ ($1 \leq k_\rho \leq d_\rho$) such that for every ρ and every k_ρ, each subrepresentation $V_\rho^{k_\rho}\colon H \to \mathbf{U}(E_\rho^{k_\rho})$ is equivalent to the irreducible representation $M_\rho\colon H \to \mathbf{U}(\mathbb{C}^{n_\rho})$.

Furthermore, each space $E_\rho^{k_\rho}$ has an H-steerable orthonormal basis with respect to an irreducible representation equivalent to M_ρ (the functions specified by the column vectors Y_{ρ,k_ρ}).

The union of these H-steerable families for all $\rho \in R(h)$ and all k_ρ is an H-steerable Hilbert basis of $\mathrm{L}^2(X)$.

As similar result is presented in Lang and Weiler [31] and in Cesa Lang and Weiler [7].

We now consider several examples.

Example 7.2. Let $H = \mathbf{SO}(2)$ and $X = S^1 \approx \mathbf{SO}(2)$. In this case $R = \mathbb{Z}$, all irreducible representations are one-dimensional and of the form $z \mapsto e^{in\theta} z$, and the characters are given by $\chi_n(e^{i\theta}) = e^{in\theta}$.

Given a function $f \in \mathrm{L}^2(S^1)$ we have (with $\varphi = \alpha - \theta$)

$$\begin{aligned}\pi_n(f)(e^{i\alpha}) &= \int_H \overline{\chi_n(e^{i\theta})} f((e^{i\theta})^{-1} e^{i\alpha})\, d\lambda(h) \\ &= \int_{-\pi}^{\pi} e^{-in\theta} f(e^{-i\theta} e^{i\alpha})\, \frac{d\theta}{2\pi} \\ &= \frac{1}{2\pi} e^{-in\alpha} \int_{-\pi}^{\pi} e^{in\varphi} f(e^{i\varphi})\, d\varphi \\ &= \frac{1}{2\pi} e^{i(-n)\alpha} \int_{-\pi}^{\pi} e^{-i(-n)\varphi} f(e^{i\varphi})\, d\varphi = e^{-in\alpha} c_{-n},\end{aligned}$$

where

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-in\varphi} f(e^{i\varphi})\, d\varphi$$

is the nth Fourier coefficient of f.

Thus

$$\pi_n(f)(e^{i\alpha}) = e^{-in\alpha}c_{-n}. \qquad \text{(str1)}$$

The index n is flipped to $-n$ due to the fact that the projection operator uses $\overline{\chi_\rho(h)}$.

The space E_n is one-dimensional and has the function

$$Y_n(e^{i\alpha}) = e^{-in\alpha} \qquad \text{(str2)}$$

as a basis. It is steerable since

$$Y_n(e^{i(\alpha-\theta)}) = e^{-in(\alpha-\theta)} = e^{in\theta}e^{-in\alpha} = e^{in\theta}Y_n(e^{i\alpha}),$$

and $\chi_n(e^{i\theta}) = e^{in\theta}$ is a character.

Example 7.3. Let $H = \mathbf{SO}(2)$ and $X = \mathbb{R}^2$. In this case, again $R = \mathbb{Z}$, all irreducible representations are one-dimensional and the characters are of the form $\chi_n(e^{i\theta}) = e^{in\theta}$.

Given any function $f \in \mathrm{L}^2(\mathbb{R}^2)$ we have

$$\begin{aligned}\pi_n(f)(x) &= \int_H \overline{\chi_n(e^{i\theta})} f(R_\theta^{-1}x)\, d\lambda(h) \\ &= \int_{-\pi}^{\pi} e^{-in\theta} f((R_{-\theta})x)\, \frac{d\theta}{2\pi},\end{aligned}$$

where R_θ is the rotation matrix

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

This time E_n is the Hilbert sum of countably many subspaces of dimension 1.

Let us compute $f_n(R_\varphi x)$ where $f_n(x) = \pi_n(f)(x)$.

We have

$$\begin{aligned}
f_n(R_\varphi x) &= \int_{-\pi}^{\pi} e^{-in\theta} f(R_{-\theta} R_\varphi x)\,\frac{d\theta}{2\pi} \\
&= \int_{-\pi}^{\pi} e^{-in\theta} f(R_{-(\theta-\varphi)} x)\,\frac{d\theta}{2\pi} \\
&= \int_{-\pi}^{\pi} e^{-in(\psi+\varphi)} f((R_{-\psi})x)\,\frac{d\psi}{2\pi} \\
&= e^{-in\varphi} \int_{-\pi}^{\pi} e^{-in\psi} f((R_{-\psi})x)\,\frac{d\psi}{2\pi} = e^{-in\varphi} f_n(x).
\end{aligned}$$

In summary,

$$f_n(R_\varphi x) = e^{-in\varphi} f_n(x). \tag{str3}$$

For $x = re_1$, $r \in \mathbb{R}_+$, where $e_1 = (1, 0)$, we get

$$f_n(rR_\varphi e_1) = e^{-in\varphi} f_n^{\mathrm{rad}}(r),$$

with

$$f_n^{\mathrm{rad}}(r) = f_n(re_1) = \int_{-\pi}^{\pi} e^{-in\theta} f(r(R_{-\theta})e_1) \frac{d\theta}{2\pi}. \tag{str4}$$

The function f_n^{rad} is called a *radial function*. It is a function defined on $\mathbb{R}_+$.

We see that in polar coordinates (r, φ),

$$f_n(r, \varphi) = e^{-in\varphi} f_n^{\mathrm{rad}}(r). \tag{str5}$$

Thus we are reduced to finding a Hilbert basis of $\mathrm{L}^2(\mathbb{R}_+)$.

There are many candidates but the Hilbert basis involving the *Hermite functions* is particularly elegant.

These are the functions

$$\psi_m(x) = e^{-\frac{x^2}{2}} H_m(x), \tag{str6}$$

where the $H_m(x)$ are *Hermite polynomials*.

The functions ψ_m are also a Hilbert basis of $\mathrm{L}^2(\mathbb{R})$; see Sansone [36], Chapter IV, Section 7 and Folland [18], Chapter 6, Section 6.4.

The Hermite polynomials are real polynomials given by the equations

$$H_m(x) = (-1)^m e^{x^2} \frac{d^m}{dx^m} e^{-x^2}. \tag{str7}$$

They are also defined by the recurrence relations

$$\begin{aligned} H_{n+2}(x) &= 2xH_{n+1}(x) - 2(n+1)H_n(x) \\ H_1(x) &= 2x \\ H_0(x) &= 1. \end{aligned}$$

From these equations the following explicit formula can be derived:

$$H_m(x) = \sum_{k=0}^{\lfloor \frac{m}{2} \rfloor} (-1)^k \frac{m(m-1)\cdots(m-2k+1)}{k!} (2x)^{m-2k};$$

see Sansone [36], Chapter IV, Section 2 and Folland [18], Chapter 6, Section 6.4.

The first six Hermite polynomials are

$$\begin{aligned} H_0(x) &= 1 & H_1(x) &= 2x \\ H_2(x) &= 4x^2 - 2 & H_3(x) &= 8x^2 - 12x \\ H_4(x) &= 16x^4 - 48x^2 + 12 & H_5(x) &= 32x^5 - 160x^3 + 120x. \end{aligned}$$

The Hermite polynomials are orthogonal with respect to the inner product

$$\langle f, g \rangle = \int_{\mathbb{R}} e^{-x^2} f(x) g(x) \, dx$$

and so the functions ψ_m are orthogonal with respect to the usual inner product on $\mathrm{L}^2(\mathbb{R})$.

They are not orthonormal because

$$\int_{-\infty}^{\infty} H_m^2(x)e^{-x^2}\,dx = \sqrt{\pi}\,2^m m!.$$

See Sansone [36], Chapter IV, Section 2.

The purpose of the term $e^{-\frac{x^2}{2}}$ is to insure that the functions ψ_m are square integrable over $\mathbb{R}$.

The Hermite polynomials are discussed quite extensively in Sansone [36], Chapter IV, Sections 2-5 and 7 and in Folland [18], Chapter 6, Section 6.4.

Then the functions

$$Y_{m,n}(r,\varphi) = e^{-in\varphi} e^{-\frac{r^2}{2}} H_m(r), \;\; m \geq 0, \qquad \text{(str8)}$$

form a steerable Hilbert basis of E_n ($n \in \mathbb{Z}$).

Indeed, we see immediately that

$$Y_{m,n}(r,\varphi-\theta) = e^{in\theta} Y_{m,n}(r,\varphi).$$

This case was also investigated by Weiler and Cesa [31] in a more informal fashion.

Example 7.4. Let $H = \mathbf{SO}(2)$ and $X = \mathrm{L}^2(\mathbf{SE}(2))$.

The action of $\mathbf{SO}(2)$ on $\mathrm{L}^2(\mathbf{SE}(2))$ is the left regular action $\mathbf{R}^{\mathbf{SO}(2)\to \mathrm{L}^2(\mathbf{SE}(2))}$ given by

$$\mathbf{R}^{\mathbf{SO}(2)\to \mathrm{L}^2(\mathbf{SE}(2))}_{R_\varphi}(f)(x,\psi) = f(R_{-\varphi}x, \psi - \varphi),$$
$$f \in \mathrm{L}^2(\mathbf{SE}(2)),\ x \in \mathbb{R}^2,\ R_\varphi \in \mathbf{SO}(2).$$

In this case, again $R = \mathbb{Z}$, all irreducible representations are one-dimensional and the characters are of the form $\chi_n(e^{i\varphi}) = e^{in\varphi}$.

Given any function $f \in \mathrm{L}^2(\mathbf{SE}(2))$ we have

$$\pi_n(f)(x,\psi) = \int_{-\pi}^{\pi} e^{-in\varphi} f(R_{-\varphi}x, \psi - \varphi)\, \frac{d\varphi}{2\pi},$$

where R_φ is the rotation matrix

$$R_\varphi = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

The space E_n is the Hilbert sum of countably many subspaces of dimension 1.

Write $f_n(x, \psi) = \pi_n(f)(x, \psi)$.

If we let $\varphi = \psi + \varphi_1$, so that $\psi - \varphi = -\varphi_1$, we obtain

$$\begin{aligned} f_n(x, \psi) &= \int_{-\pi}^{\pi} e^{-in\varphi} f(R_{-\varphi}x, \psi - \varphi)\, \frac{d\varphi}{2\pi} \\ &= \int_{-\pi}^{\pi} e^{-in(\psi+\varphi_1)} f(R_{-(\psi+\varphi_1)}x, -\varphi_1)\, \frac{d\varphi_1}{2\pi} \\ &= e^{-in\psi} \int_{-\pi}^{\pi} e^{-in\varphi_1} f(R_{-\varphi_1} R_{-\psi}x, -\varphi_1)\, \frac{d\varphi_1}{2\pi} \\ &= e^{-in\psi} \int_{-\pi}^{\pi} e^{in\varphi_1} f(R_{\varphi_1} R_{-\psi}x, \varphi_1)\, \frac{d\varphi_1}{2\pi}. \end{aligned}$$

In summary, we proved that

$$f_n(x,\psi) = e^{-in\psi}\int_{-\pi}^{\pi} e^{in\varphi_1} f(R_{\varphi_1}R_{-\psi}x,\varphi_1)\,\frac{d\varphi_1}{2\pi}. \quad \text{(str9)}$$

As a consequence, we have

$$f_n(R_\alpha x,\psi+\alpha) = e^{-in\alpha}f_n(x,\psi). \quad \text{(str10)}$$

For $x = re_1$, $r\in\mathbb{R}_+$, with $e_1 = (1,0)$, from (str9) we get

$$f_n(rR_\alpha e_1,\theta) = e^{-in\theta}f_n^{\mathrm{rad}}(r,\theta-\alpha), \quad \text{(str11)}$$

with

$$f_n^{\mathrm{rad}}(r,\psi) = f_n(re_1,\psi) = \int_{-\pi}^{\pi} e^{in\varphi} f(rR_\varphi R_{-\psi}e_1,\varphi)\,\frac{d\varphi}{2\pi}. \quad \text{(str12)}$$

In polar coordinates (r, α),

$$f_n((r,\alpha),\theta) = e^{-in\theta} f_n^{\mathrm{rad}}(r, \theta - \alpha). \qquad \text{(str13)}$$

Observe that since in polar coordinates the effect of a rotation $R_{-\varphi}$ is to transform (r, α) to $(r, \alpha - \varphi)$, we have

$$\begin{aligned} f_n((r,\alpha-\varphi),\theta-\varphi) &= e^{-in(\theta-\varphi)} f_n^{\mathrm{rad}}(r, \theta-\varphi-(\alpha-\varphi)) \\ &= e^{in\varphi} e^{-in\theta} f_n^{\mathrm{rad}}(r, \theta-\alpha) \\ &= e^{in\varphi} f_n((r,\alpha),\theta), \end{aligned}$$

confirming that the functions f_n are steerable.

The functions f_n^{rad} belong to $\mathrm{L}^2(\mathbb{R}_+ \times \mathbf{SO}(2))$.

Since the functions $e^{-\frac{r^2}{2}} H_m(r)$ form a Hilbert basis of $\mathrm{L}^2(\mathbb{R}_+)$ and the functions $e^{-ik\psi}$ form a Hilbert basis of $\mathrm{L}^2(\mathbf{SO}(2))$, it can be shown that the family of functions

$$e^{-\frac{r^2}{2}} H_m(r) e^{-ik\psi}$$

form a Hilbert basis of $\mathrm{L}^2(\mathbb{R}_+ \times \mathbf{SO}(2))$; see Lang [32], Chapter XVII, Problem 9.

At first glance it is not obvious that the functions $f_n^{\mathrm{rad}}(r, \psi)$ yield all the functions in the Hilbert basis of $\mathrm{L}^2(\mathbb{R}_+ \times \mathbf{SO}(2))$.

In fact they do.

By (str13) and the above reasoning, the functions

$$e^{-in\theta}e^{ik(\theta-\alpha)}e^{-\frac{r^2}{2}}H_m(r)$$

for n fixed form a Hilbert basis of E_n, and thus the functions

$$\begin{aligned} Y_{k,m,n}((r,\alpha),\theta) &= e^{-in\theta}e^{ik(\theta-\alpha)}e^{-\frac{r^2}{2}}H_m(r) \\ &= e^{-i(n-k)\theta}e^{-ik\alpha}e^{-\frac{r^2}{2}}H_m(r) \qquad \text{(str14)} \end{aligned}$$

form a steerable basis of $\mathrm{L}^2(\mathbf{SE}(2))$, with $n, k \in \mathbb{Z}$ and $m \geq 0$.

In Section **??** it will be more convenient to change the index k to $n-k$, in which case the term $e^{-i(n-k)\theta}e^{-ik\alpha}$ becomes

$$e^{-ik\theta}e^{-i(n-k)\alpha} = e^{-in\alpha}e^{-ik(\theta-\alpha)},$$

and so we also have the steerable basis of functions

$$e^{-in\alpha}e^{-ik(\theta-\alpha)}e^{-\frac{r^2}{2}}H_m(r), \quad n,k \in \mathbb{Z},\ m \geq 0. \quad \text{(str15)}$$

Example 7.5. Let H be any compact group and let $X = G$ with G acting on itself by left multiplication.

Since the M_ρ are (irreducible) representations of H we have $M_\rho(s^{-1}t) = M_\rho(s^{-1})M_\rho(t) = M_\rho(s)^* M_\rho(t)$, so *the jth column* $(1/n_\rho)m^{(\rho)}_{*j}(s^{-1}t)$ of the matrix $M_\rho(s^{-1}t)$ can be expressed as

$$\begin{aligned}(1/n_\rho)m^{(\rho)}_{*j}(s^{-1}t) &= M_\rho(s)^*(1/n_\rho)m^{(\rho)}_{*j}(t)\\ &= (\overline{M_\rho(s)})^\top(1/n_\rho)m^{(\rho)}_{*j}(t),\end{aligned}$$

and so

$$\overline{m^{(\rho)}_{*j}}(s^{-1}t) = (M_\rho(s))^\top\overline{m^{(\rho)}_{*j}}(t). \qquad \text{(str16)}$$

Since the family of functions

$$\left(\frac{1}{\sqrt{n_\rho}} m_{ij}^{(\rho)}\right)_{1 \leq i,j \leq n_\rho,\, \rho \in R(H)}$$

is a Hilbert basis of $\mathrm{L}^2(G)$, it follows that according to Definition 7.2,

$$(\overline{m_{1j}}, \ldots, \overline{m_{n_\rho, j}})$$

forms a steerable basis of $\overline{\mathfrak{l}_j^{(\rho)}}$ for $j = 1, \ldots, n_\rho$, using the notation of Section 4.2.

Note that in terms of the notation used in Theorem 7.1, $d_\rho = n_\rho$.

Recall that by Peter–Weyl I, $\mathrm{L}^2(H)$ is the Hilbert sum of minimal two-sided ideals $\mathfrak{a}_\rho$ isomorphic to the matrix algebra $\mathrm{M}_{n_\rho}(\mathbb{C})$, and $\mathfrak{a}_\rho$ is expressed as the finite Hilbert sum of n_ρ minimal left ideals $\mathfrak{l}_j^{(\rho)}$.

Observe that we can also obtain the above result by considering the left regular representation $V = \mathbf{R}$ of G.

As noted just after Definition 4.9, the projection π_ρ^V maps $\mathrm{L}^2(G)$ onto $\overline{\mathfrak{a}_\rho}$, so the functions $(\overline{m_{1j}}, \ldots, \overline{m_{n_\rho,j}})$ are indeed in $\overline{\mathfrak{a}_\rho}$ and form a basis of $\overline{\mathfrak{l}_j^{(\rho)}}$, the jth column of $\overline{M}_\rho$.

Thus Equation (steer6) holds.

Example 7.6. In this example we describe a method generalizing the method of Example 7.4 to decompose $\mathrm{L}^2(\mathbf{SE}(n))$ using the representation V and the projections

$$\begin{aligned}(\pi_\rho^V(f))(x, h_1) &= n_\rho \int_{\mathbf{SO}(n)} \overline{\chi_\rho(h)} f(h^{-1}x, h^{-1}h_1)\, d\lambda(h)\\ &= \int_{\mathbf{SO}(n)} \overline{u_\rho(h)} f(h^{-1}x, h^{-1}h_1)\, d\lambda(h),\end{aligned}$$

with $(x, h_1) \in \mathbf{SE}(n)$, for all $f \in \mathrm{L}^2(\mathbf{SE}(n))$ and all $\rho \in R(\mathbf{SO}(n))$.

Write $f^\rho(x, h_1) = (\pi_\rho^V(f))(x, h_1)$.

Since $\mathbf{SO}(n)$ is unimodular (because it is compact), with $h = h_1 h_2$, we have

$$\begin{aligned} f^\rho(x, h_1) &= \int_{\mathbf{SO}(n)} \overline{u_\rho(h)} f(h^{-1}x, h^{-1}h_1)\, d\lambda(h), \\ &= \int_{\mathbf{SO}(n)} \overline{u_\rho(h_1 h_2)} f(h_2^{-1}h_1^{-1}x, h_2^{-1})\, d\lambda(h_2) \\ &= \int_{\mathbf{SO}(n)} \overline{u_\rho(h_1 h_2^{-1})} f(h_2 h_1^{-1}x, h_2)\, d\lambda(h_2). \end{aligned} \tag{$*_{13}$}$$

Recall that $u_\rho = m_{11}^{(\rho)} + \cdots + m_{n_\rho n_\rho}^{(\rho)}$, which is n_ρ times the trace of the matrix M_ρ corresponding to the irreducible representation of $\mathbf{SO}(n)$ indexed by ρ.

Since $M_\rho = (1/n_\rho)\left(m^{(\rho)}_{ij}\right)$ and it is a representation, because $(1/n_\rho)m^{(\rho)}_{ii}(h_1h_3)$ is the (i,i)-entry in the matrix $M_\rho(h_1h_3)$, it is equal to the inner product of the ith row of $M_\rho(h_1)$ by the ith column of $M_\rho(h_3)$, so

$$(1/n_\rho)m^{(\rho)}_{ii}(h_1h_3) = \sum_{j=1}^{n_\rho}(1/n_\rho)m^{(\rho)}_{ij}(h_1)(1/n_\rho)m^{(\rho)}_{ji}(h_3),$$

and by multiplying both sides by n_ρ we get

$$u_\rho(h_1h_3) = \sum_{i=1}^{n_\rho} m^{(\rho)}_{ii}(h_1h_3) = (1/n_\rho)\sum_{i,j=1}^{n_\rho} m^{(\rho)}_{ij}(h_1)m^{(\rho)}_{ji}(h_3). \tag{$*_{14}$}$$

The calculations in $(*_{14})$ and $(*_{13})$ imply that

$$\begin{aligned} f^\rho(x, h_1) &= \int_{\mathbf{SO}(n)} \overline{u_\rho(h_1 h_2^{-1})} f(h_2 h_1^{-1} x, h_2)\, d\lambda(h_2) \\ &= n_\rho \sum_{i,j=1}^{n_\rho} \frac{1}{n_\rho} \overline{m_{ij}^{(\rho)}(h_1)} \\ &\quad \int_{\mathbf{SO}(n)} \frac{1}{n_\rho} \overline{m_{ji}^{(\rho)}(h_2^{-1})} f(h_2 h_1^{-1} x, h_2)\, d\lambda(h_2) \\ &= n_\rho \sum_{i,j=1}^{n_\rho} \frac{1}{n_\rho} \overline{m_{ij}^{(\rho)}(h_1)} \\ &\quad \int_{\mathbf{SO}(n)} \frac{1}{n_\rho} m_{ij}^{(\rho)}(h_2) f(h_2 h_1^{-1} x, h_2)\, d\lambda(h_2). \end{aligned} \tag{$*_{15}$}$$

Using the fact that if A and B are any two $n \times n$ matrices, then

$$\sum_{i,j=1}^{n} a_{ij} b_{ij} = \mathrm{tr}(AB^\top)$$

and observing that the matrix whose entries are the terms

$$\int_{\mathbf{SO}(n)} \frac{1}{n_\rho} m_{ij}^{(\rho)}(h_2) f(h_2 h_1^{-1} x, h_2)\, d\lambda(h_2)$$

is the matrix

$$\begin{aligned}
&\int_{\mathbf{SO}(n)} M_\rho(h_2) f(h_2 h_1^{-1} x, h_2)\, d\lambda(h_2) \\
&= \int_{\mathbf{SO}(n)} \left(\left(\overline{M_\rho(h_2)} \right)^* \right)^\top f(h_2 h_1^{-1} x, h_2)\, d\lambda(h_2),
\end{aligned}$$

we obtain

$$\begin{aligned}
f^\rho(x, h_1) = n_\rho \operatorname{tr}\Bigg(& \overline{M_\rho(h_1)} \\
& \int_{\mathbf{SO}(n)} \left(\overline{M_\rho(h_2)} \right)^* f(h_2 h_1^{-1} x, h_2)\, d\lambda(h_2) \Bigg).
\end{aligned} \tag{$*_{16}$}$$

Observe that this is the generalization of (str9).

Also, $\left(\overline{M_\rho(h_2)}\right)^* = \left(M_\rho(h_2)\right)^\top$.

We also define $f_\rho^{\mathrm{rad}}\colon \mathbb{R}^d \to \mathrm{M}_{n_\rho}(\mathbb{C})$ by

$$\begin{aligned} f_\rho^{\mathrm{rad}}(x) &= \int_{\mathbf{SO}(n)} \left(\overline{M_\rho(h_2)}\right)^* f(h_2 x, h_2)\, d\lambda(h_2) \\ &= \int_{\mathbf{SO}(n)} M_\rho(h_2)^\top f(h_2 x, h_2)\, d\lambda(h_2), \qquad (f_\rho^{\mathrm{rad}}) \end{aligned}$$

and so we have

$$f^\rho(x, h_1) = n_\rho \operatorname{tr}\left(\overline{M_\rho(h_1)}\, f_\rho^{\mathrm{rad}}(h_1^{-1} x)\right). \qquad (f^\rho)$$

Observe that

$$\begin{aligned} f^{\rho}(hx, hh_1) &= n_\rho \, \mathrm{tr}\Big(\overline{M_\rho(hh_1)}\, f_\rho^{\mathrm{rad}}((hh_1)^{-1}hx)\Big) \\ &= n_\rho \mathrm{tr}\Big(\overline{M_\rho(h)}\,\overline{M_\rho(h_1)}\, f_\rho^{\mathrm{rad}}(h_1^{-1}x)\Big), \end{aligned} \tag{str17}$$

which expresses steerability with respect to $\mathbf{SO}(n)$.

Using Lang [32], Chapter XVII, Problem 9, since the family of functions $e^{-\frac{x^2}{2}} H_m(x)$ is a Hilbert basis of $\mathrm{L}^2(\mathbb{R})$, the family of functions

$$\begin{aligned} & e^{-\frac{x_1^2}{2}} H_{k_1}(x_1) \cdots e^{-\frac{x_n^2}{2}} H_{k_n}(x_n) \\ & = e^{-\frac{\|x\|_2^2}{2}} H_{k_1}(x_1) \cdots H_{k_n}(x_n), \quad k_1, \ldots, k_n \geq 0, \end{aligned}$$

with $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, is a Hilbert basis of $\mathrm{L}^2(\mathbb{R}^n)$.

For $f \in \mathrm{L}^2(\mathbf{SE}(n))$ given by

$$f(x, h_2) = e^{-\|x\|^2/2} H_{k_1}(x_1) \cdots H_{k_n}(x_n)\, \overline{m_{\ell k}^{(\rho)}(h_2)},$$

we find that $f_\rho^{\mathrm{rad}}(x)$ is the $n_\rho \times n_\rho$-matrix whose (ℓ, k) entry is $e^{-\|x\|^2/2} H_{k_1}(x_1) \cdots H_{k_n}(x_n)$, and all other entries are 0, which implies that

$$f^\rho(x, h_1) = \overline{m_{k\ell}^{(\rho)}(h_1)}\, e^{-\|x\|^2/2} H_{k_1}((h_1^{-1}x)_1) \cdots H_{k_n}((h_1^{-1}x)_n)$$

belongs to the subspace E_ρ, the projection of $\mathrm{L}^2(\mathbf{SE}(n))$ by π_ρ^V.

The Hilbert space $\mathrm{L}^2(\mathbf{SE}(n))$ is isomorphic to $\mathrm{L}^2(\mathbf{SO}(n) \times \mathbb{R}^n)$, and since by Peter-Weyl I, the Hilbert space $\mathrm{L}^2(\mathbf{SO}(n))$ is the Hilbert sum of the minimal two-sided ideals $\mathfrak{a}_\rho$ which have the n_ρ^2 functions $\overline{m_{k\ell}^{(\rho)}}$ as an orthogonal basis, we conclude that the family of functions

$$\left(\overline{m_{k\ell}^{(\rho)}(h_1)}\, e^{-\frac{\|x\|^2}{2}} H_{k_1}((h_1^{-1}x)_1) \cdots H_{k_n}((h_1^{-1}x)_n)\right) \Bigg| \\ \rho \in R(\mathbf{SO}(n)),\ 1 \leq k, \ell \leq n_\rho,\ k_1, \ldots, k_n \geq 0, \tag{str18}$$

with $h_1 \in \mathbf{SO}(n)$ and $x \in \mathbb{R}^n$, is an $\mathbf{SO}(n)$-steerable Hilbert basis of $\mathrm{L}^2(\mathbf{SE}(n))$.

More precisely, for any fixed $\rho \in R(\mathbf{SO}(n)), 1 \leq \ell \leq n_\rho$, $k_1, \ldots, k_n \geq 0$, if we write $\mathbf{k} = (k_1, \ldots, k_n)$, by (str16), the column vector $Y^{\rho}_{\ell,\mathbf{k}}(h_1, x)$ of dimension n_ρ with

$$Y^{\rho}_{k,\ell,\mathbf{k}}(h_1, x) = \overline{m^{(\rho)}_{k\ell}(h_1)}\, e^{-\frac{\|x\|^2}{2}} \\ H_{k_1}((h_1^{-1}x)_1) \cdots H_{k_n}((h_1^{-1}x)_n), \\ 1 \leq k \leq n_\rho,$$

satisfies the streerability equation

$$Y^{\rho}_{\ell,\mathbf{k}}(h^{-1}h_1, h^{-1}x) = (M_\rho(h))^\top Y^{\rho}_{\ell,\mathbf{k}}(h_1, x). \qquad \text{(str19)}$$

If $n = 2$, then $R(\mathbf{SO}(2)) = \mathbb{Z}$, $m_\ell(\theta) = e^{i\ell\theta}$, so we find that the family

$$\Big(e^{-i\ell\theta} e^{-\frac{x^2+y^2}{2}} H_{k_1}(x\cos\theta + y\sin\theta) \\ H_{k_2}(-x\sin\theta + y\cos\theta)\Big)_{\ell\in\mathbb{Z},\, k_1,k_2\geq 0} \tag{str20}$$

is steerable basis of $\mathrm{L}^2(\mathbf{SE}(2))$.

If $n = 3$, then $\overline{R(\mathbf{SO}(3))} = \mathbb{N}$, $\rho = \ell$, $n_\rho = 2\ell + 1$, the functions $\sqrt{2\ell+1}\, w_{jk}^{(\ell)}(R)$ $(R \in \mathbf{SO}(3))$ of Section 5.15 (see also Section 5.10) from a Hilbert basis of $\mathrm{L}^2(\mathbf{SO}(3))$, so we find that the family

$$\Big(\sqrt{2\ell+1}\,\overline{w_{jk}^{(\ell)}(R)}\, e^{-\frac{x_1^2+x_2^2+x_3^2}{2}} \\ H_{k_1}((R^{-1}x)_1) H_{k_2}((R^{-1}x)_2) H_{k_3}((R^{-1}x)_3)\Big) \Big| \\ \ell \in \mathbb{N},\ -\ell \leq j, k \leq \ell,\ k_1, k_2, k_3 \geq 0, \qquad \text{(str21)}$$

with $R \in \mathbf{SO}(3)$ and $x \in \mathbb{R}^3$, is steerable basis of $\mathrm{L}^2(\mathbf{SE}(3))$.

For fixed $\ell \in \mathbb{N}, -\ell \leq k \leq \ell, k_1.k_2, k_3 \geq 0$, if we write $\mathbf{k} = (k_1, k_2, k_3)$ and if $Y^{\ell}_{k,\mathbf{k}}(R, x)$ is the column vector given by

$$\begin{aligned} Y^{\ell}_{j,k,\mathbf{k}}(R, x) = \sqrt{2\ell+1}\,&\overline{w^{(\ell)}_{jk}(R)}e^{-\frac{x_1^2+x_2^2+x_3^2}{2}} \\ &H_{k_1}((R^{-1}x)_1)H_{k_2}((R^{-1}x)_2)H_{k_3}((R^{-1}x)_3), \\ &-\ell \leq j \leq \ell, \end{aligned}$$

then we have

$$Y^{\ell}_{k,\mathbf{k}}(Q^{-1}R, Q^{-1}x) = (w^{(\ell)}(Q))^{\top} Y^{\ell}_{k,\mathbf{k}}(R, x). \quad \text{(str22)}$$

We can also express the matrices $w^{(\ell)}(R)$ in terms of the Euler angles and the Wigner d-matrices as in Section 5.15; see the Remark just after Proposition 5.20.

In Section 7.1 we noticed that the functions $\widehat{f}$ are *vector-valued* functions from $\mathbb{R}^d$ to the codomain $\mathrm{M}_L(\mathbb{C})$ and that the group $G = \mathbb{R}^d \rtimes H$ acts on their domain $\mathbb{R}^d$, whereas the group H acts on their codomain $\mathrm{M}_L(\mathbb{C})$ in terms of the representation Σ.

Experience has shown that the design of efficient convolution neural networks (CNN) is greatly facilitated if they operate on functions having the properties of the $\widehat{f}$ listed above. Such functions are known as *feature fields*.