197010 Groups and Physics

Roy Meshulam

July 13, 2022

1 Introduction to Representation Theory

Let G be a finite group and V a finite dimensional complex vector space.

A *representation* of G on V is a homomorphism $\rho : G \to GL(V)$. The *degree* of ρ is $\dim V$. By identifying $GL_1(\mathbb{C})$ with $\mathbb{C}^*$, we can view a 1-dimensional representation as a homomorphism $\chi : G \to \mathbb{C}^*$.

Examples.
1. The trivial representation 1_G *of* G *on* V *is given by* $1_G(g)(v) = v$ *for all* $g \in G$ *and* $v \in V$.
2. The sign representation *of the symmetric group* S_n *is the homomorphism* $sgn : S_n \to \mathbb{C}^*$, *where* $sgn(\pi)$ *is the sign of* π.
3. Let $C_n = \langle x \rangle$ *be the cyclic group of order* n *with a generator* x. *For* $k \in \mathbb{Z}_n$ *let* $\chi_k : C_n \to \mathbb{C}^*$ *be given by* $\chi_k(x^\ell) = \exp(\frac{2\pi i k\ell}{n})$.
4. Let $G = S_n$, $V = \{x = (x_1, \ldots, x_n) \in \mathbb{C}^n : \sum_{i=1}^n x_i = 0\}$. *The* natural representation *of* S_n *on* V *is given by* $\rho(\sigma)(x) = (x_{\sigma^{-1}(1)}, \ldots, x_{\sigma^{-1}(n)})$.
5. Let G *act on the left on a finite set* X. *Let* $V = span\{e_x : x \in X\}$ *be the complex vector spanned by the basis* $\{e_x : x \in X\}$. *The* regular representation $\mathrm{reg}_{G,X}$ *of* G *is given on the basis elements by* $\mathrm{reg}_{G,X}(e_x) = e_{gx}$. *Thus* $\mathrm{reg}_{G,X}(\sum_{x\in X} a_x e_x) = \sum_{x \in X} a_x e_{gx}$. *For the case* $X = G$ *and the action of* G *is by left multiplication, the regular representation is denoted by* reg_G.
6. Let $G = D_n = \langle s, r : s^2 = r^n = 1, srs = r^{-1} \rangle$ *be the dihedral group. Let* $\omega = \exp\left(\frac{2\pi i}{n}\right)$. *Let* ψ_0, ψ_1 *be the* 1*-dimensional representations of* D_n *given by*

	r^ℓ	sr^ℓ
ψ_0	1	1
ψ_1	1	-1

For $0 \le k \le n-1$ *let* ρ_k *be the* 2*-dimensional representation given by*

$$\rho_k(r^\ell) = \begin{bmatrix} \omega^{k\ell} & 0 \\ 0 & \omega^{-k\ell} \end{bmatrix} \quad , \quad \rho_k(sr^\ell) = \begin{bmatrix} 0 & \omega^{-k\ell} \\ \omega^{k\ell} & 0 \end{bmatrix}.$$

Odd n: *Then* $\{\psi_0, \psi_1\} \cup \{\rho_k : 1 \le k \le \frac{n-1}{2}\}$ *are all irreducible representations of* D_n.
Even n: *Let* ψ_2, ψ_3 *be the* 1*-dimensional representations of* D_n *given by*

	r^ℓ	sr^ℓ
ψ_2	$(-1)^\ell$	$(-1)^\ell$
ψ_3	$(-1)^\ell$	$(-1)^{\ell+1}$

Then $\{\psi_i : 0 \leq i \leq 3\} \cup \{\rho_k : 1 \leq k \leq \frac{n}{2} - 1\}$ *are all irreducible representations of* D_n.

1.1 Basic Properties

Let (V, ρ) a representation of G. A subspace $W \subset V$ is *invariant* if $\rho(g)W = W$ for all $g \in G$. In this case (W, ρ) is a representation of G.

Claim 1.1. *For any representation* (V, ρ) *of* G *there exists an inner product* $\langle \cdot, \cdot \rangle$ *on* V *such that* $\langle \rho(g)u, \rho(g)v \rangle = \langle u, v \rangle$ *for all* $u, v \in V$ *and* $g \in G$.

Proof. Let $(\cdot, \cdot)$ be an arbitrary inner product on V, and let $\langle u, v \rangle = \sum_{g \in G} (\rho(g)u, \rho(g)v)$. Then for any $h \in G$

$$\begin{aligned} \langle \rho(h)u, \rho(h)v \rangle &= \sum_{g \in G} (\rho(g)\rho(h)u, \rho(g)\rho(h)v) \\ &= \sum_{g \in G} (\rho(gh)u, \rho(gh)v) = \langle u, v \rangle. \end{aligned}$$

□

Claim 1.2. *Let* (V, ρ) *a representation of* G *and let* $U \subset V$ *be an invariant subspace. Then there exists an invariant subspace* $W \subset V$ *such that* $V = U \oplus W$.

Proof. Let
$$W = U^\perp = \{w \in V : \langle w, u \rangle = 0 \text{ for all } u \in U\}.$$
Then $U \oplus W = V$ and W is invariant. Indeed, if $w \in W$, then $\langle gw, u \rangle = \langle w, g^{-1}u \rangle = 0$ for any $u \in U$.

□

A representation (V, ρ) is *irreducible* if it does not have nontrivial (i.e. different from 0 and V) invariant subspaces.

Corollary 1.3. *Any representation* (V, ρ) *is a direct sum* $V = V_1 \oplus \cdots \oplus V_k$ *of irreducible representations.*

For two representation (V_1, ρ_1), (V_2, ρ_2) of G, let $\mathrm{Hom}(V_1, V_2)$ denote the space of linear maps from V_1 to V_2, and let

$$\mathrm{Hom}_G(V_1, V_2) = \{T \in \mathrm{Hom}(V_1, V_2) : T\rho_1(g) = \rho_2(g)T\}$$

denote the space of linear G-maps from V_1 to V_2. Thus $T \in \mathrm{Hom}_G(V_1, V_2)$ iff the diagram

$$\begin{array}{ccc} V_1 & \xrightarrow{T} & V_2 \\ \rho_1(g)\downarrow & & \downarrow\rho_2(g) \\ V_1 & \xrightarrow{T} & V_2 \end{array}$$

is commuting for every $g \in G$. The representations (V_1, ρ_1), (V_2, ρ_2) are *isomorphic*, denoted by $\rho_1 \cong \rho_2$, if there exists an isomorphism $T \in \mathrm{Hom}_G(V_1, V_2)$.

Proposition 1.4 (Schur's Lemma)**.** *For two irreducible representations* (V_1, ρ_1), (V_2, ρ_2) *of* G*:*

$$\dim \mathrm{Hom}_G(V_1, V_2) = \begin{cases} 1 & \rho_1 \cong \rho_2, \\ 0 & \rho_1 \ncong \rho_2. \end{cases}$$

Proof. Let $T \in \mathrm{Hom}_G(V_1, V_2)$. Then $\ker T$ is a G-invariant subspace of V_1. Indeed, if $u_1 \in \ker T$ then $T\rho_1(g)u_1 = \rho_2(g)Tu_1 = 0$, hence $\rho_1(g)u_1 \in \ker T$. similarly, $T(V_1)$ is G-invariant subspace of V_2, as $\rho_2(g)T(V_1) = T\rho_1(V_1) = T(V_1)$. Hence, if $T \neq 0$ then $T(V_1) \neq 0$ and therefore $T(V_1) = V_2$. Moreover, $\ker T \neq V_1$ and hence $\ker T = 0$, i.e. T is an isomorphism. Thus $\rho_1 \ncong \rho_2$ implies that $\mathrm{Hom}_G(V_1, V_2) = 0$. On the other hand, suppose that $\rho_1 \cong \rho_2$. We may then assume that $V_1 = V_2 = U$ and $\rho_1 = \rho_2 = \rho$. Let $T \in \mathrm{Hom}_G(U, U)$ and let $\lambda \in \mathbb{C}$ be an eigenvalue of T. Then $\ker(T - \lambda I) \neq 0$ and therefore $\ker(T - \lambda I) = U$, i.e. $T = \lambda I$.

□

Corollary 1.5. *If G is abelian and (V, ρ) is an irreducible representation of G, then* $\dim V = 1$.

Proof. Fix $h \in G$. Then for any $g \in G$

$$\rho(h)\rho(g) = \rho(hg) = \rho(gh) = \rho(g)\rho(h).$$

Therefore $\rho(h) \in \mathrm{Hom}_G(V, V)$. It follows by Schur's Lemma that there exist a $\lambda_h \in \mathbb{C}$ such that $\rho(h) = \lambda_h I$. Therefore any 1-dimensional of V is invariant. Hence $\dim V = 1$.

□

1.2 Operations on Representations

Let (V_1, ρ_1), (V_2, ρ_2) be representations of G. The *direct sum* representation is $(V_1 \oplus V_2, \rho_1 \oplus \rho_2)$, where $(\rho_1 \oplus \rho_2)(g)(v_1, v_2) = (\rho_1(g)v_1, \rho_2(g)v_2)$. The *tensor product* representation $(V_1 \otimes V_2, \rho_1 \otimes \rho_2)$ where the action is given by

$$(\rho_1 \otimes \rho_2)(g)(v_1 \otimes v_2) = \rho_1(g)v_1 \otimes \rho_2(g)v_2$$

on decomposable elements $v_1 \otimes v_2$, and extended by linearity to the whole of $V_1 \otimes V_2$. The *dual* of (V, ρ) is (V^*, ρ^*) where for $g \in G$, $\phi \in V^*$ and $v \in V$ we define $\rho^*(g)\phi(v) = \phi(\rho(g^{-1})v)$. The Hom representation of (V_1, ρ_1), (V_2, ρ_2) is $(\mathrm{Hom}(V_1, V_2), \mathrm{Hom}(\rho_1, \rho_2)$ where for $g \in G$, $\phi \in V^*$ and $v_1 \in V_2$ we define

$$\mathrm{Hom}(\rho_1, \rho_2)(g)\phi(v_1) = \rho_2(g)\phi(\rho_1(g^{-1})v_1).$$

Remarks.
1. If ρ_2 is the trivial representation of G on $V_2 = \mathbb{C}$ then

$$(\mathrm{Hom}(V_1, V_2), \mathrm{Hom}(\rho_1, \rho_2)) \cong (V_1^*, \rho_1^*).$$

2. define $T : V_1^* \otimes V_2 \to \mathrm{Hom}(V_1, V_2)$ by $T(\phi \otimes v_2)(v_1) = \phi(v_1)v_2$. Then

$$\begin{aligned}
\mathrm{Hom}(\rho_1, \rho_2)(g)T(\phi \otimes v_2)(v_1) &= \rho_2(g)T(\phi \otimes v_2)(\rho_1(g^{-1})v_1) \\
&= \rho_2(g)\left(\phi(\rho_1(g^{-1})v_1)v_2\right) \\
&= \phi(\rho_1(g^{-1})v_1)\rho_2(g)v_2 \\
&= \rho_1^*(g)\phi(v_1) \otimes \rho_2(g)v_2 \\
&= T(\rho_1^*(g)\phi \otimes \rho_2(g)v_2)v_1 \\
&= T(\rho_1^* \otimes \rho_2(g))(\phi \otimes v_2)(v_1).
\end{aligned}$$

It follows that the diagram

$$\begin{array}{ccc}
V_1^* \otimes V_2 & \xrightarrow{T} & \mathrm{Hom}(V_1, V_2) \\
{\scriptstyle \rho_1^* \otimes \rho_2(g)} \downarrow & & \downarrow {\scriptstyle \mathrm{Hom}(\rho_1, \rho_2)(g)} \\
V_1^* \otimes V_2 & \xrightarrow{T} & \mathrm{Hom}(V_1, V_2)
\end{array}$$

for all $g \in G$ and therefore $(V_1^* \otimes V_2, \rho_1^* \otimes \rho_2) \cong (\mathrm{Hom}(V_1, V_2), \mathrm{Hom}(\rho_1, \rho_2))$.

1.3 Characters

Let $L(G)$ denote the space of complex valued functions on G. Let $L_c(G) \subset L(G)$ denote the subspace of *class functions* on G, i.e. all $f \in L(G)$ such that $f(hgh^{-1}) = f(g)$ for all $g, h \in G$. The *character* of a representation (V, ρ) is the function $\chi_\rho \in L(G)$ given by $\chi_\rho(g) = \mathrm{tr}\rho(g)$. Here are some basic properties of χ_ρ.
1. $\chi_\rho(1) = \dim V = \deg \rho$.
2. $\chi_\rho \in L_c(G)$. Indeed,

$$\chi_\rho(hgh^{-1}) = \mathrm{tr}(\rho(hgh^{-1})) = \mathrm{tr}(\rho(h)\rho(g)\rho(h)^{-1}) = \mathrm{tr}(\rho(g)) = \chi_\rho(g).$$

3. $\chi_\rho(g^{-1}) = \overline{\chi_\rho(g)}$. Let $\langle \cdot, \cdot \rangle$ be a G-invariant inner product on V, and let $v_1, \ldots, v_n$

be an orthonormal basis of V. Then $\chi_\rho(g) = \sum_{i=1}^n \langle \rho(g)v_i, v_i\rangle$. It follows that

$$\begin{aligned}\chi_\rho(g^{-1}) &= \sum_{i=1}^n \langle \rho(g^{-1})v_i, v_i\rangle = \sum_{i=1}^n \langle v_i, \rho(g)v_i\rangle \\ &= \sum_{i=1}^n \overline{\langle \rho(g)v_i, v_i\rangle} = \overline{\chi_\rho(g)}.\end{aligned}$$

Examples.
1. $\chi_\rho(1) = \dim V = \deg \rho$.
2. If ρ *is* 1*-dimensional, then* $\chi_\rho(g) = \rho(g)$.
3. Let G *act on* X *and let* ρ *be the permutation representation on* $V = span\{e_x : x \in X\}$, $\rho(g)e_x = e_{gx}$. *Then*

$$\chi_\rho(g) = |\mathrm{Fix}(g)| = |\{x \in X : gx = x\}|.$$

Let (V_i, ρ_i), $i = 1, 2$ be two representations.

Claim 1.6. *(i)* $\chi_{\rho_1 \oplus \rho_2}(g) = \chi_{\rho_1}(g) + \chi_{\rho_2}(g)$. *(ii)* $\chi_{\rho_1 \otimes \rho_2}(g) = \chi_{\rho_1}(g) \cdot \chi_{\rho_2}(g)$. *(iii)* $\chi_{\rho^*}(g) = \overline{\chi_\rho(g)}$. *(iv)* $\chi_{\mathrm{Hom}(\rho_1, \rho_2)}(g) = \overline{\chi_{\rho_1}(g)} \cdot \chi_{\rho_2}(g)$.

Proof. (i) and (ii) follow from the facts that if A_1, A_2 are two square matrices then $\mathrm{tr}(A_1 \oplus A_2) = \mathrm{tr}(A_1) + \mathrm{tr}(A_2)$ and $\mathrm{tr}(A_1 \otimes A_2) = \mathrm{tr}(A_1) \cdot \mathrm{tr}(A_2)$. For (iii), Let $B = \{v_1, \ldots, v_n\}$ be an orthonormal basis of V with respect to $\langle \cdot, \cdot \rangle$. Let $B^* = \{\phi_1, \ldots, \phi_n\}$ be a basis of V^* dual to B, i.e. $\phi_i(v_j) = \delta_{ij}$. Let $A = (a_{ij})$ be the matrix representing $\rho(g^{-1})$ with respect to B, i.e. $\rho(g^{-1})v_j = \sum_{k=1}^n a_{kj}v_k$. Then A^t is the matrix representing $\rho^*(g)$ with respect to B^*. Indeed,

$$\begin{aligned}\rho^*(g)\phi_i(v_j) &= \phi_i(\rho(g^{-1})v_j) \\ &= \phi_i(\sum_{k=1}^n a_{kj}v_k) = \sum_{k=1}^n a_{kj}\phi_i(v_k) \\ &= \sum_{k=1}^n a_{kj}\delta_{ik} = a_{ij} = \sum_{k=1}^n a_{ik}\phi_k(v_j).\end{aligned} \tag{1}$$

Combining (1) with the fact that $\rho(g) \in U(n)$ we obtain

$$\chi_{\rho^*}(g) = \mathrm{tr}\rho(g^{-1}) = \mathrm{tr}\rho(g)^{-1} = \mathrm{tr}\rho(g)^* = \overline{\chi_\rho(g)}.$$

$\square$

Let V be a finite dimensional complex vector space and let $U \subset V$ be a subspace. A linear map $P \in \mathrm{Hom}(V, U)$ is a *projection* onto U if $Pu = u$ for all $u \in U$ and $P^2 = P$. Recall the following

Claim 1.7. *If* $P : U \to V$ *is a projection then* $trP = \dim U$.

Proof. Note that $\ker P \cap U = 0$ and $\ker P + U = V$, hence $\ker P \oplus U = V$. Let $B_1 = \{u_1, \ldots, u_k\}$ be a basis of U, and let $B_2 = \{v_1, \ldots, v_\ell\}$ be a basis of $\ker P$. Then the matrix representing P according to the basis $B_1 \cup B_2$ of V is $M = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix}$, hence $\mathrm{tr}P = \mathrm{tr}M = k = \dim U$.

□

Let (V, ρ) be a representation of G. In the sequel we will often abbreviate gv for $\rho(g)v$. Let $V^G = \{v \in V : gv = v \text{ for all } g \in G\}$ be the *fixed subspace* for the action of G.

Claim 1.8. *The mapping $P : V \to V$ given by $Pv = \frac{1}{|G|}\sum_{g\in G} gv$ is a projection of V onto V^G.*

Corollary 1.9.

$$\dim V^G = trP = \frac{1}{|G|}\sum_{g\in G} tr\rho(g) = \frac{1}{|G|}\sum_{g\in G} \chi_\rho(g).$$

Define an inner product on $L(G)$ by

$$(\phi, \psi) = \frac{1}{|G|}\sum_{g\in G} \phi(g)\overline{\psi(g)}.$$

Claim 1.10. *Let (V_1, ρ_1), (V_2, ρ_2) be two irreducible representations of G. Then*

$$(\chi_{\rho_1}, \chi_{\rho_2}) = \begin{cases} 1 & \rho_1 \cong \rho_2, \\ 0 & \rho_1 \ncong \rho_2. \end{cases} \tag{2}$$

Proof. First note that the fixed subspace $\mathrm{Hom}(V_1, V_2)^G$ of the representation $\mathrm{Hom}(\rho_1, \rho_2)$ on $\mathrm{Hom}(V_1, V_2)$ is $\mathrm{Hom}_G(V_1, V_2)$. Using Schur's lemma we compute

$$\begin{aligned}
\overline{(\chi_{\rho_1}, \chi_{\rho_2})} = (\chi_{\rho_2}, \chi_{\rho_1}) &= \frac{1}{|G|}\sum_{g\in G} \chi_{\rho_2}(g)\overline{\chi_{\rho_1}(g)} \\
&= \frac{1}{|G|}\sum_{g\in G} \chi_{\rho_1^*\otimes\rho_2}(g) = \frac{1}{|G|}\sum_{g\in G} \chi_{\mathrm{Hom}(\rho_1,\rho_2)}(g) \\
&= \dim \mathrm{Hom}(V_1, V_2)^G = \dim \mathrm{Hom}_G(V_1, V_2) = \begin{cases} 1 & \rho_1 \cong \rho_2, \\ 0 & \rho_1 \ncong \rho_2. \end{cases}
\end{aligned}$$

□

Let $\{(W_i, \rho_i)\}_{i=1}^t$ be the irreducible representations of G and write $\chi_i = \chi_{\rho_i}$. We have shown that $\{\chi_i\}_{i=1}^t$ is an orthonormal, and in particular independent family in $L_c(G)$.

Corollary 1.11. *Let (V, ρ) be a representation of G, and let*

$$V = \oplus_{j=1}^{m} U_j \tag{3}$$

be a decomposition of V into irreducible representations. Then

(i)

$$\ell_i := |\{1 \leq j \leq m : (U_j, \rho) \cong (W_i, \rho_i)\}| = (\chi_i, \chi_\rho).$$

In particular, ℓ_i is independent of the particular decomposition (3).

(ii) $(V, \rho) \cong (V', \rho')$ *iff* $\chi_\rho = \chi_{\rho'}$.

(iii) (V, ρ) *is irreducible iff* $(\chi_\rho, \chi_\rho) = 1$.

(iv) For $1 \leq i \leq t$, the space $L_i = \oplus\{U_j : (U_j, \rho) \cong (W_i, \rho_i)\}$ is independent of the decomposition (3). A projection of V onto L_i is given by

$$P_i = \frac{\chi_i(1)}{|G|} \sum_{x \in G} \overline{\chi_i(x)} \rho(x).$$

L_i is called the isotypic component *of V corresponding to the W_i.*

1.4 The Fourier Transform

Let $f \in L(G)$. The *F*ourier Transform $\widehat{f}(\rho)$ of f at a representation (V, ρ) of G is given by

$$\widehat{f}(\rho) = \sum_{x \in G} f(x) \rho(x) \in \mathrm{End}(V).$$

Let $\{(V_i, \rho_i) : 1 \leq i \leq t\}$ be the irreps of G and let $d_i = \dim V_i$.

Claim 1.12 (Fourier inversion formula)**.** *For any $x \in G$*

$$f(x) = \frac{1}{|G|} \sum_{i=1}^{t} d_i tr\left(\widehat{f}(\rho_i) \rho_i(x^{-1})\right). \tag{4}$$

Proof.

$$
\begin{aligned}
&\frac{1}{|G|}\sum_{i=1}^{t} d_i \mathrm{tr}\left(\widehat{f}(\rho_i)\rho_i(x^{-1})\right)\\
&= \frac{1}{|G|}\sum_{i=1}^{t} d_i \mathrm{tr}\left(\sum_{y\in G} f(y)\rho_i(y)\cdot\rho_i(x^{-1})\right)\\
&= \frac{1}{|G|}\sum_{i=1}^{t} d_i \mathrm{tr}\left(\sum_{y\in G} f(y)\rho_i(yx^{-1})\right)\\
&= \frac{1}{|G|}\sum_{i=1}^{t} d_i \left(\sum_{y\in G} f(y)\chi_i(yx^{-1})\right) \qquad (5)\\
&= \frac{1}{|G|}\sum_{y\in G} f(y)\left(\sum_{i=1}^{t} d_i\chi_i(yx^{-1})\right)\\
&= \frac{1}{|G|}\sum_{y\in G} f(y)\left(\sum_{i=1}^{t} \overline{\chi_i(1)}\chi_i(yx^{-1})\right)\\
&= \frac{1}{|G|}\sum_{y\in G} f(y)|G|\delta_{1,yx^{-1}} = f(x).
\end{aligned}
$$

We identify $L(G)$ with the group algebra $\mathbb{C}[G]$ by associating to $f \in L(G)$, the element $\sum_{x\in G} f(x)x$. Under this identification, the convolution of $f, g \in L(G)$ is given by $f * g(x) = \sum_{y\in G} f(y)g(y^{-1}x)$ is mapped to the product

$$
\left(\sum_{x\in G} f(x)x\right)\cdot\left(\sum_{y\in G} f(y)y\right).
$$

Claim 1.13. *For $f, g \in L(G)$ and a representation (V, ρ)*

$$
\widehat{f * g}(\rho) = \widehat{f}(\rho)\cdot\widehat{g}(\rho). \qquad (6)
$$

Proof.

$$
\begin{aligned}
\widehat{f * g}(\rho) &= \sum_{x\in G} f * g(x)\rho(x)\\
&= \sum_{x\in G}\left(\sum_{y\in G} f(y)g(y^{-1}x)\right)\rho(x)\\
&= \sum_{y\in G} f(y)\rho(y)\left(\sum_{x\in G} g(y^{-1}x)\rho(y^{-1}x)\right)\\
&= \left(\sum_{y\in G} f(y)\rho(y)\right)\left(\sum_{x\in G} g(x)\rho(x)\right) = \widehat{f}(\rho)\cdot\widehat{g}(\rho).
\end{aligned}
$$

□

Define an inner product on $\prod_{i=1}^t \mathrm{End}(V_i)$ by

$$\big((S_i)_{i=1}^t, (T_i)_{i=1}^t\big) = \sum_{i=1}^t d_i \mathrm{tr}(S_i T_i^*).$$

The Fourier Map $F : \mathbb{C}[G] \to \prod_{i=1}^t \mathrm{End}(V_i)$ is given by

$$F\left(\sum_{x\in G} f(x)x\right) = \left(\widehat{f}(\rho_1), \ldots, \widehat{f}(\rho_t)\right).$$

Claim 1.14. *(i) The Fourier map is an isomorphism of algebras. (ii) Any $\phi, \psi \in L(G)$ satisfy the Parseval identity*

$$|G|^2(\phi, \psi) = (F(\phi), F(\psi)). \tag{7}$$

Proof. (i) Let $G : \prod_{i=1}^t \mathrm{End}(V_i) \to \mathbb{C}[G]$ be given by

$$G(T_1, \ldots, T_t) = \frac{1}{|G|}\sum_{x\in G}\left(\sum_{i=1}^t d_i \mathrm{tr}(T_i \rho(x^{-1}))\right) x.$$

Then $GF = \mathit{Id}$ by Claim 1.12. As

$$\dim \mathbb{C}[G] = |G| = \sum_{i=1}^t d_i^2 = \dim \prod_{i=1}^t \mathrm{End}(V_i)$$

Together with Claim 1.13 it follows that F is an isomorphism of algebras.
(ii)

$$\begin{aligned}
(F(\phi), F(\psi)) &= \sum_{i=1}^t d_i \mathrm{tr}\left(\widehat{\phi}(\rho_i)\widehat{\psi}(\rho_i)^*\right) \\
&= \sum_{i=1}^t d_i \mathrm{tr}\left(\left(\sum_{x\in G}\phi(x)\rho_i(x)\right)\cdot\left(\sum_{y\in G}\overline{\psi(y)}\rho_i(y)^*\right)\right) \\
&= \sum_{i=1}^t d_i \sum_{x,y\in G}\phi(x)\overline{\psi(y)}\mathrm{tr}\left(\rho_i(x)\rho_i(y)^*\right) \\
&= \sum_{x,y\in G}\phi(x)\overline{\psi(y)}\left(\sum_{i=1}^t d_i\chi_i(xy^{-1})\right) \\
&= |G|\sum_{x,y\in G}\phi(x)\overline{\psi(y)}\delta_{xy} = |G|^2(\phi,\psi).
\end{aligned}$$

□

1.5 Discrete Vector Bundles

Let X be a finite set. A *discrete vector bundle* over the base space X is a family of pairwise disjoint complex vector spaces $\{E_x\}_{x\in X}$. The *total space* of the bundle is $E = \cup_{x\in X} E_x$. Define $\pi : E \to X$ by $\pi(e) = x$ if $e \in E_x$. We will often refer to E as a vector bundle over X. A *section* of E is a map $s : X \to E$ such that $\pi(s(x)) = x$, i.e. s sends x to the fiber E_x. Let $\Gamma(E)$ denote the set of all sections of E. The natural pointwise addition $(s_1 + s_2)(x) = s_1(x) + s_2(x)$, and multiplication by scalar $(cs)(x) = cs(x)$, make $\Gamma(E)$ into a complex vector space of dimension $\dim \Gamma(X) = \sum_{x\in X} \dim E_x$.

Suppose now that G is a finite group that acts both on X and on E. We say that E is a *G-vector bundle* over X if the following holds:

- $\pi(g(e)) = g(\pi(e))$ for any $e \in E$ and $g \in G$. In other words, g maps E_x into E_{gx}.
- The map $g : E_x \to E_{gx}$ is linear.

Given a G-vector bundle E over X, define a representation ρ of G on $\Gamma(E)$ as follows. For $s \in \Gamma(E)$ and $x \in X$ let

$$\rho(g)(s)(x) = g\left(s(g^{-1}x)\right).$$

Example 1.15. *Let G act on a finite set X. For $x \in X$ let $E_x = \{x\} \times \mathbb{C}$ be a fixed one dimensional space, and let G acts on E by $g(x, \lambda) = (gx, \lambda)$. In other words, the composition $\mathbb{C} \to E_x \to E_{gx} \to \mathbb{C}$ is the identity map. In this case, the representation ρ on $\Gamma(E)$ is isomorphic to the permutation representation of G on X.*

Example 1.16. *Let X be a finite set in $\mathbb{R}^d$, and let G be a subgroup of the orthogonal group $O(d)$ that permutes the elements of X. For $x \in X$, let $E_x = \{x\} \times \mathbb{C}^d$, and define the action of G on E by $g(x, u) = (gx, gu)$. This type of action will later occur in our discussion of molecular vibrations. Figure 2(a) depicts a section $s \in \Gamma(E)$, and Figure 2(b) depicts $\rho(g)s(x)$, where $g \in U(2)$ is the $\frac{\pi}{2}$-rotation.*

We next compute the character of ρ. For $x \in X$ let $\dim E_x = d_x$, and let $u_{x1}, \ldots, u_{xd_x}$ be an arbitrary basis of E_x. For $1 \le i \le d_x$ define a section $s_{xi} \in \Gamma(X)$ by

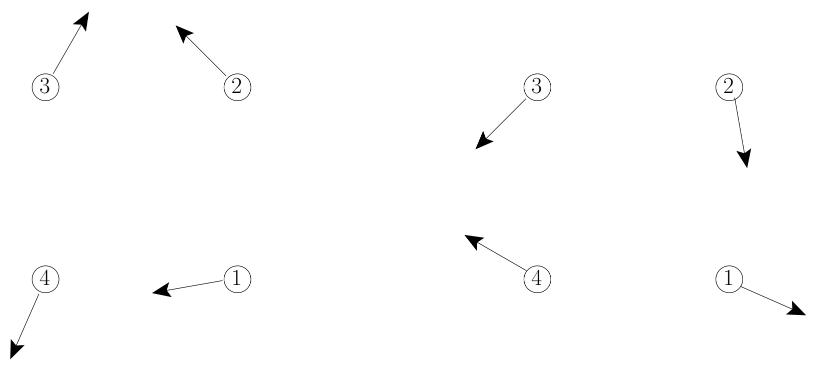
$$s_{xi}(y) = \delta_{x,y} u_{xi} = \begin{cases} u_{xi} & y = x, \\ 0 & \text{otherwise.} \end{cases}$$

The following statement is straightforward.

Claim 1.17. *$\{s_{xi} : x \in X, 1 \le i \le d_x\}$ is a basis of $\Gamma(X)$.*

We next compute the character χ_ρ. Let $g \in G$, $x \in X$ and $1 \le i \le d_x$. Then $gu_{xi} \in E_{gx}$ and hence there exist coefficients $\{\alpha_{xi,j}(g)\}_{j=1}^{d_x}$ such that

$$gu_{xi} = \sum_j \alpha_{xi,j}(g) u_{gx,j}.$$



(a) $s(x)$ (b) $\rho(g)s(x) = gs(g^{-1}x)$

Figure 1: The action of G on $\Gamma(E)$

Claim 1.18.

$$\rho(g)(s_{xi}) = \sum_j \alpha_{xi,j}(g) s_{gx,j}. \tag{8}$$

Proof. Let $y \in x$. Then

$$\begin{aligned} \rho(g)(s_{xi})(y) &= g(s_{xi}(g^{-1}y)) \\ &= \delta_{g^{-1}y,x} g u_{xi} = \delta_{g^{-1}y,x} \sum_j \alpha_{xi,j}(g) u_{gx,j} \\ &= \sum_j \alpha_{xi,j}(g) \left(\delta_{gx,y} u_{gx,j}\right) \\ &= \sum_j \alpha_{xi,j}(g) s_{gx,j}(y). \end{aligned}$$

□

Corollary 1.19.

$$\chi_\rho(g) = \sum_{\{x \in X : gx = x\}} tr\left[g : E_x \to E_x\right]. \tag{9}$$

Proof. By Claim 1.18

$$\chi_\rho(g) = \sum_{\{x \in X : gx = x\}} \sum_i \alpha_{xi,i}(g) = \sum_{\{x \in X : gx = x\}} \mathrm{tr}\left[g : E_x \to E_x\right].$$

□

1.6 Induced Representations

One of the main tools for constructing representations is by using induction from representations of subgroups. In this subsection we only consider the finite group case, but we note that while technically more involved, the method of induction works for infinite groups as well.

Let G be a finite group and let $H < G$ be a subgroup. Let $\rho : G \to GL(V)$ be a representation of G. The restriction of ρ to H is denoted by $\mathrm{Res}_H^G(\rho)$. We now introduce the dual operation of inducing a representation of G from a representation of H.

Let $\lambda : H \to GL(W)$ be a complex finite dimensional representation of H. Define a vector bundle $W_\lambda := G \times_H W$ over the coset space $X = G/H$ as follows. For $g \in G$ let $[g] = gH$ be the corresponding coset. Let $\sim$ be the equivalence relation on $G \times W$ given by $(g, w) \sim (gh, \lambda(h^{-1})w)$ for all $(g, h, w) \in G \times H \times W$. Let W_λ be the quotient space $(G \times W)/\sim$. Denote by $[x, w]$ the equivalence class of $(x, w) \in G \times W$. The projection map $[g, w] \to [g] = gH$ defines a vector bundle over G/H. Define an action of G on G/H and on W_λ by $g(xH) = gxH$. Define an action of G on $G \times W$ by $g(x, w) = (gx, w)$. Clearly, if $(x_1, w_1) \sim (x_2, w_2)$, then $g(x_1, w_1) \sim g(x_2, w_2)$. Thus we get an action of G on W_λ. Note that this is *not* a trivial action on the fibers. Indeed, let $R = \{g_1, \ldots, g_m\}$ be cosets representatives for H in G, and let define a vector bundle $E = R \times W$. Let $F : E \to W_\lambda$ be the isomorphism $F(g_i, w) = [g_i, w]$. The action of G on W_λ gives an action of G on E by $g(e) = F^{-1}g(F(e))$ for $g \in G$. Let's describe this action explicitly. Suppose that $e = (g_i, w)$. Then $gg_i = g_jh$ for some $1 \le j \le m$ and $h \in H$. Then

$$\begin{aligned} g(g_i, w) &= g(e) = F^{-1}g(F(e)) \\ &= F^{-1}g[g_i, w] = F^{-1}[gg_i, w] \\ &= F^{-1}[g_jh, w] = F^{-1}[g_j, \lambda(h)w] \\ &= (g_j, \lambda(h)w) = (g_j, \lambda(g_j^{-1}gg_i)w). \end{aligned} \tag{10}$$

The *induced representation* $\mathrm{Ind}_H^G\lambda$ is the representation ρ of G on $\Gamma(W_\lambda)$, given by $\rho(g)s(x) = gs(g^{-1}x)$ for $x \in X = G/H$. It will often be convenient to work with the following isomorphic version of $\Gamma(W_\lambda)$. Let $C(G, W)$ denote the space of all functions from G to W. Let

$$C_{H,\lambda}(G, W) = \{\phi \in C(G, W) : \phi(xh) = \rho(h^{-1})\phi(x) \text{ for all } x \in G, h \in H\}.$$

Let η denote the representation of G on V given by $\eta(g)\phi(x) = \phi(g^{-1}x)$, for $g, x \in G$. Define $T : C_{H,\lambda}(G, W) \to \Gamma(W_\lambda)$ by $T\phi([x]) = [x, \phi(x)]$. Note that T is well defined, i.e. if $[x_1] = [x_2]$ then $x_2 = x_1h$ for some $h \in H$ and thus

$$(x_2, \phi(x_2)) = (x_1h, \phi(x_1h)) = (x_1h, \lambda(h^{-1})\phi(x_1))$$

and hence $(x_2, \phi(x_2)) \sim (x_1, \phi(x_1))$.

Claim 1.20. *T is an isomorphism and the following diagram commutes:*

$$\begin{array}{ccc} C_{H,\lambda}(G,W) & \xrightarrow{T} & \Gamma(W_\lambda) \\ {\scriptstyle \eta(g)}\downarrow & & \downarrow{\scriptstyle \rho(g)} \\ C_{H,\lambda}(G,W) & \xrightarrow{T} & \Gamma(W_\lambda) \end{array}$$

Proof. Let $\phi \in C_{H,\lambda}(G,W)$ and let $x \in G$. Then

$$\begin{aligned} \rho(g)T\phi([x]) &= g(T\phi(g^{-1}[x])) \\ &= g[g^{-1}x, \phi(g^{-1}x)] = [x, \phi(g^{-1}x)] \\ &= [x, \eta(g)\phi(x)] = T\left(\eta(g)\phi\right)[x] \end{aligned}$$

□

In view of Claim 1.20 we will identify $\Gamma(W_\lambda)$ with $C_{H,\lambda}(G,W)$ and the representation η with $\rho = \mathrm{Ind}_H^G\lambda$. Note that the degree of ρ is

$$\deg \mathrm{Ind}_H^G\lambda = \dim \Gamma(W_\lambda) = \frac{|G|}{|H|} \cdot \dim W = \frac{|G|}{|H|} \cdot \deg \lambda.$$

For explicit computations, e.g. for finding actual matrix forms of the representation it is often convenient to fix a set $\{g_1, \ldots, g_m\}$ of coset representatives of H in G, i.e. $G = \cup_{i=1}^m g_i H$, where $m = \frac{|G|}{|H|}$. For $g \in G$, let $\pi(g, g_i)$ denote the unique g_j such that $gg_i \in g_j H$. For $1 \le i \le m$ and $w \in W$, let $e_{g_i,w}$ be the unique element in $C_{H,\lambda}(G,W)$ that satisfies $e_{g_i,w}(g_j) = \delta_{ij} w$. Clearly if $w_1, \ldots, w_d$ is a basis of W, then $\left\{e_{g_i,w_j} : 1 \le i \le m, 1 \le j \le d\right\}$ is a basis of $C_{H,\lambda}(G,W)$. The induced representation $\rho = \mathrm{Ind}_H^G\lambda$ of G on $C_{H,\lambda}(G,W)$ is given as follows. If $\pi(g, g_i) = g_j$, then

$$\rho(g)e_{g_i,w} = e_{g_j,\lambda(g_j^{-1}gg_i)w}.$$

Example 1.21. *Let $\lambda = \mathbf{1}$ be the trivial representation of H on $W = \mathbf{C}$. Then $\rho = \mathrm{Ind}_H^G\lambda$ is the permutation representation of G on G/H. Indeed, Let $g_1, \ldots, g_m$ be coset representatives of H. Then $\{e_{g_i,1} : 1 \le i \le m\}$ is a basis of $C_{H,\lambda}(G, \mathbf{C})$. Let $g \in G$ and let $\pi(g, g_i) = g_j$. Then*

$$\rho(g)e_{g_i,1}(g_k) = e_{g_i,1}(g^{-1}g_k) = \delta_{kj} = e_{g_j,1}(g_k).$$

Example 1.22. *Let $G = D_n = \langle s, r : s^2 = r^n = 1, srs = r^{-1}\rangle$ be the dihedral group. Let $\omega = \exp\left(\frac{2\pi i}{n}\right)$. Let $N = \langle r\rangle$ and for $k \in \mathbf{Z}_n$, let χ_k be the character of N given by $\chi_k(r^\ell) = \omega^{k\ell}$. We compute $\rho_k = \mathrm{Ind}_N^G\chi_k$. Following the general recipe as above, let $g_1 = 1, g_2 = s$ be coset representatives for N. Then $\{e_{1,1}, e_{s,1}\}$ is a basis of $C_{H,\chi_k}(G, \mathbf{C})$.*

	1	s
$e_{1,1}$	1	0
$e_{s,1}$	0	1

.

Next note that

	1	s
$\pi(r^\ell,\cdot)$	1	s
$\pi(sr^\ell,\cdot)$	s	1

.

Therefore

	$e_{1,1}$	$e_{s,1}$
$\rho(r^\ell)$	$\chi_k(r^\ell)e_{1,1}$	$\chi_k(r^{-\ell})e_{s,1}$
$\rho(sr^\ell)$	$\chi_k(r^\ell)e_{s,1}$	$\chi_k(r^{-\ell})e_{1,1}$

.

It follows that the matrices representing $\rho(g)$ *with respect to the basis* $e_{1,1}, e_{s,1}$ *are given by*

$$\rho(r^\ell) = \begin{bmatrix} \omega^{k\ell} & 0 \\ 0 & \omega^{-k\ell} \end{bmatrix} \quad , \quad \rho(sr^\ell) = \begin{bmatrix} 0 & \omega^{-k\ell} \\ \omega^{k\ell} & 0 \end{bmatrix}.$$

We next compute the character of the induced representation.

Claim 1.23. *Let* λ *be a representation of* H *and let* $\rho = \mathrm{Ind}_H^G\lambda$*. Then for* $g \in G$

$$\chi_\rho(g) = \frac{1}{|H|} \sum_{\{x\in G: x^{-1}gx\in H\}} \chi_\lambda(x^{-1}gx). \tag{11}$$

Proof. By (9)

$$\begin{aligned} \chi_\rho(g) &= \sum_{\{1\le i\le m: \pi(g,g_i)=g_i\}} \mathrm{tr}[g : E_{g_i} \to E_{g_i}] \\ &= \sum_{\{1\le i\le m: \pi(g,g_i)=g_i\}} \mathrm{tr}[\lambda(g_i^{-1}gg_i) : W \to W] \\ &= \sum_{\{1\le i\le m: \pi(g,g_i)=g_i\}} \chi_\lambda(g_i^{-1}gg_i) \\ &= \frac{1}{|H|} \sum_{\{x\in G: x^{-1}gx\in H\}} \chi_\lambda(x^{-1}gx). \end{aligned} \tag{12}$$

□

Proposition 1.24 (Frobenius Reciprocity)**.** *Let* $\eta : G \to GL(U)$ *be a representation of* G*, and let* $\lambda : H \to GL(W)$ *be a representation of* H*. Then*

$$\mathrm{Hom}_G(\eta, \mathrm{Ind}_H^G\lambda) \cong \mathrm{Hom}_H(\mathrm{Res}_H^G\eta, \lambda). \tag{13}$$

In particular, writing $\rho = \mathrm{Ind}_H^G\lambda$ *we have*

$$(\chi_\eta, \chi_\rho)_G = \left(\chi_{\eta|H}, \chi_\lambda\right)_H. \tag{14}$$

Proof. Later.

1.7 Representations of Semidirect Products

One useful application of induction concerns the representations of semidirect products. Let H be a group that acts on the left an abelian group N. The *semidirect product* $H \ltimes N$ is the group whose elements are $H \times N$, with the product given by

$$(h_1, n_1) \cdot (h_2, n_2) = (h_1 h_2, n_1 + h_1(n_2)) .$$

Encoding (h, n) by the matrix $\begin{pmatrix} h & n \\ 0 & 1 \end{pmatrix}$ the product rule in $H \ltimes N$ becomes

$$\begin{pmatrix} h_1 & n_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} h_2 & n_2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} h_1 h_2 & n_1 + h_1(n_2) \\ 0 & 1 \end{pmatrix}.$$

Furthermore, $(h, n)^{-1} = (h^{-1}, -h^{-1}(n))$ and $(h, 0)(1, n)(h, 0)^{-1} = (1, h(n))$.

Example 1.25 (The Dihedral Group)**.** *Let $S = \{1, s\}$ the cyclic group of order 2. Then S acts on $\mathbf{Z}_n$ by $s^\epsilon(k) = (-1)^\epsilon k$. The semidirect product $S \ltimes \mathbf{Z}_n$ is isomorphic to the dihedral group $D_n = \langle s, r : s^2 = r^n = 1, srs = r^{-1}\rangle$ via the map $f : (s^\epsilon, k) \to r^k \cdot s^\epsilon$. Indeed,*

$$\begin{aligned} &f\left((s^{\epsilon_1}, k_1) \cdot (s^{\epsilon_2}, k_2)\right) = f\left(s^{\epsilon_1+\epsilon_2}, k_1 + (-1)^{\epsilon_1} k_2\right) \\ &= r^{k_1+(-1)^{\epsilon_1} k_2} s^{\epsilon_1+\epsilon_2} = r^{k_1} s^{\epsilon_1} \cdot \left(s^{\epsilon_1} r^{(-1)^{\epsilon_1} k_2} s^{\epsilon_1}\right) s^{\epsilon_2} \\ &= r^{k_1} s^{\epsilon_1} \cdot r^{k_2} s^{\epsilon_2} = f(s^{\epsilon_1}, k_1) \cdot f(s^{\epsilon_2}, k_2). \end{aligned}$$

Example 1.26 (The metabelian group of order pq)**.** *Let p, q be primes numbers such that p divides $q - 1$. Let λ be a multiplicative generator of $\mathbf{Z}_q^*$. Let $r = \frac{q-1}{p}$ and let $\alpha = \lambda^r$. There is a unique nonabelian group $D_{p,q}$ of order pq. It has two generators a, b and has the following presentation*

$$D_{p,q} = \langle a, b : a^p = b^q = 1 \ , \ aba^{-1} = b^\alpha\rangle.$$

D_{pq} is isomorphic to the semidirect product $C_p \ltimes \mathbf{Z}_q$ where $C_p = \langle a\rangle$ acts on $\mathbf{Z}_q$ be $a(m) = \alpha m$. The map $f : C_p \ltimes \mathbf{Z}_q \to D_{pq}$ is given by $f(a^k, \ell) = b^\ell a^k$ is the required isomorphism.

Example 1.27 (The Finite Affine Group)**.** *Let p be prime and let $\mathbf{F}_p$ be field with p elements. Let*

$$\mathrm{Aff}(\mathbf{F}_p) = \left\{ \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} : a \in \mathbf{F}_p^*, b \in \mathbf{F}_p \right\} \subset GL_2(\mathbf{F}_p).$$

$\mathrm{Aff}(\mathbf{F}_p)$ is isomorphic to the semidirect product $\mathbf{F}_p^* \ltimes \mathbf{F}_p$, where the action of $\mathbf{F}_p^*$ on $\mathbf{F}_p$ is given by $a(b) = ab$. The map $f : \mathbf{F}_p^* \ltimes \mathbf{F}_p \to \mathrm{Aff}(\mathbf{F}_p)$ given by $(a, b) \to \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix}$ is the required isomorphism.

Example 1.28 (The Finite Heisenberg Group)**.** *Let*

$$H(p) = \left\{ \begin{bmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix} \; : \; a, b, c \in \mathbf{F}_p \right\}.$$

Let the cyclic group $C_p = \langle s : s^p = 1 \rangle$ *act on* $\mathbf{F}_p^2$ *by* $s^a(b, c) = (b, c + ab)$*. The mapping* $f : C_p \ltimes \mathbf{F}_p^2 \to H(p)$ *given by*

$$f\left(s^a, (b, c)\right) = \begin{bmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix}$$

is an isomorphism.

We now discuss the representation theory of semidirect products. Let the group H act on the abelian group N and let $G = H \ltimes N$ be the associated semidirect product. Let $\widehat{N}$ denote the character group of N. Then H acts on $\widehat{N}$ by

$$h(\phi)(n) = \phi(h^{-1}(n)).$$

Let $\phi_1, \ldots, \phi_m$ be representatives of the orbits of $\widehat{N}$ under H, an let $K_i = \mathrm{Stab}_H(\phi_i) < H$ be the stabilizer of ϕ_i. Let $\lambda_{ij} : K_i \to GL(W_{ij})$ be all irreducible representations of K_i, where $\dim W_{ij} = d_{ij}$ and $1 \le j \le r_i$. Let $\lambda_{ij} \ltimes \phi_i : K_i \ltimes N \to \mathrm{GL}(W_i)$ be given by $(\lambda_{ij} \ltimes \phi_i)(k,n) = \phi_i(n)\lambda_{ij}(k)$. Let $\rho_{ij} = \mathrm{Ind}_{K_i \ltimes N}^{G} (\lambda_{ij} \ltimes \phi_i)$.

Proposition 1.29.

(i) All ρ_{ij} are distinct and irreducible.

(ii) $\mathrm{Irr}(G) = \{\rho_{ij} : 1 \le i \le m \ , \ 1 \le j \le r_i\}$.

Proof. (i) Let $(h_0, n_0) \in H \ltimes N$. Fix $\phi \in \widehat{N}$, $K = \mathrm{Stab}_H(\phi)$ and an irreducible representation $\lambda : K \to W$. Let $\rho = \mathrm{Ind}_{K \ltimes N}^{G} (\lambda \ltimes \phi)$. Then for $(h_0, n_0) \in G$

$$\begin{aligned}
\chi_\rho(h_0,n_0) &= \frac{1}{|K|\cdot|N|} \sum_{(h,n)^{-1}(h_0,n_0)(h,n)\in K\ltimes N} \chi_{\lambda\ltimes\phi}\left((h,n)^{-1}(h_0,n_0)(h,n)\right) \\
&= \frac{1}{|K|\cdot|N|} \sum_{(h,n)^{-1}(h_0,n_0)(h,n)\in K\ltimes N} \chi_{\lambda\ltimes\phi}\left(h^{-1}h_0h, -h^{-1}(n) + h^{-1}(n_0) + h^{-1}h_0(n)\right) \\
&= \frac{1}{|K|\cdot|N|} \sum_{(h,n)^{-1}(h_0,n_0)(h,n)\in K\ltimes N} \chi_{\lambda\ltimes\phi}\left(h^{-1}h_0h, h^{-1}(n_0)\right) \\
&= \frac{1}{|K|\cdot|N|} \sum_{n\in N} \sum_{h^{-1}h_0h\in K} \phi(h^{-1}(n_0))\chi_\lambda\left(h^{-1}h_0h\right) \\
&= \frac{1}{|K|} \sum_{h^{-1}h_0h\in K} h(\phi)(n_0)\chi_\lambda\left(h^{-1}h_0h\right).
\end{aligned}$$

Suppose now that $\phi_1, \phi_2 \in \widehat{N}$ are two characters of N, such that either $\phi_1 = \phi_2$ or that ϕ_1 and ϕ_2 are in different orbits of H. For $i = 1, 2$ let $K_i = \mathrm{Stab}_H\phi_i$ and let $\lambda_i : K_i \to GL(W_i)$ be an irreducible representation of K_i. Let $\rho_i = \mathrm{Ind}_{K_i \ltimes N}^{G} (\lambda_i \ltimes \phi_i)$. Then

$$\begin{aligned}
|G|\left(\chi_{\rho_1}, \chi_{\rho_2}\right) &= \sum_{(h_0,n_0)} \chi_{\rho_1}(h_0,n_0)\overline{\chi_{\rho_2}(h_0,n_0)} \\
&= \frac{1}{|K_1|\cdot|K_2|} \sum_{h_0} \sum_{\substack{h_1^{-1}h_0h_1\in K_1 \\ h_2^{-1}h_0h_2\in K_2}} \chi_{\lambda_1}(h_1^{-1}h_0h_1)\overline{\chi_{\lambda_2}(h_2^{-1}h_0h_2)} \sum_{n_0} h_1(\phi_1)(n_0)\overline{h_2(\phi_2)(n_0)} \\
&= \frac{|N|}{|K_1|\cdot|K_2|} \sum_{h_0} \sum_{\substack{h_1^{-1}h_0h_1\in K_1 \\ h_2^{-1}h_0h_2\in K_2}} \delta_{h_1(\phi_1),h_2(\phi_2)}\chi_{\lambda_1}(h_1^{-1}h_0h_1)\overline{\chi_{\lambda_2}(h_2^{-1}h_0h_2)}
\end{aligned}$$

If $\phi_1 \neq \phi_2$ then by assumption ϕ_1 are ϕ_2 are in different orbits of H, and hence $\delta_{h_1(\phi_1),h_2(\phi_2)} = 0$ for all $h_1, h_2 \in H$ and therefore $(\chi_{\rho_1}, \chi_{\rho_2}) = 0$. Otherwise $\phi_1 = \phi_2 = \phi$ and $K_1 = K_2 = K$. Therefore

$$\begin{aligned}
&|G|\,(\chi_{\rho_1}, \chi_{\rho_2}) = \\
&= \frac{|N|}{|K|^2} \sum_{h_0} \sum_{\substack{h_1^{-1}h_0h_1 \in K \\ h_2^{-1}h_0h_2 \in K}} \delta_{h_1(\phi),h_2(\phi)} \chi_{\lambda_1}(h_1^{-1}h_0h_1)\overline{\chi_{\lambda_2}(h_2^{-1}h_0h_2)} \\
&= \frac{|N|}{|K|^2} \sum_{h_0} \sum_{h^{-1}h_0h \in K} \sum_{k \in K} \chi_{\lambda_1}(h^{-1}h_0h)\overline{\chi_{\lambda_2}((hk)^{-1}h_0(hk))} \\
&= \frac{|N|}{|K|} \sum_{h_0} \sum_{h^{-1}h_0h \in K} \chi_{\lambda_1}(h^{-1}h_0h)\overline{\chi_{\lambda_2}(h^{-1}h_0h)} \\
&= \frac{|N|}{|K|} \sum_{\{(h_0,h) \in H \times H : h^{-1}h_0h \in K\}} \chi_{\lambda_1}(h^{-1}h_0h)\overline{\chi_{\lambda_2}(h^{-1}h_0h)} \\
&= \frac{|N|}{|K|} \sum_{(k,h) \in K \times H} \chi_{\lambda_1}(k)\overline{\chi_{\lambda_2}(k)} \\
&= |G|(\chi_{\lambda_1}, \chi_{\lambda_2})_K = |G|\delta_{\lambda_1,\lambda_2}.
\end{aligned}$$

(ii) The degree of $\mathrm{Ind}_{K_i \ltimes N}^{G}(\lambda_{ij} \ltimes \phi_i)$ is $\frac{|H|}{|K_i|}d_{ij}$, and

$$\begin{aligned}
\sum_{i=1}^{m} \sum_{j=1}^{r_i} \left(\frac{|H|}{|K_i|}d_{ij}\right)^2 &= |H|^2 \sum_{i=1}^{m} \frac{1}{|K_i|^2} \sum_{j=1}^{r_i} d_{ij}^2 = |H|^2 \sum_{i=1}^{m} \frac{1}{|K_i|} \\
&= |H|^2 \sum_{i=1}^{m} \frac{|H(\phi_i)|}{|H|} = |H| \cdot |N| = |G|.
\end{aligned}$$

□

Example 1.30 (Representations of the finite affine group). *Recall that* $G = \mathrm{Aff}(\mathbf{F}_p) = H \ltimes N$ *where* $H = \mathbf{F}_p^*$, $N = \mathbf{F}_p$ *and* H *act on* N *by* $a(b) = ab$ *for* $a \in H, b \in N$. *Let* $\widehat{N} = \{\phi_k\}_{k=0}^{p-1}$ *be the character group of* N, *where* $\phi_k(n) = w^{kn}$, $w = \exp(\frac{2\pi i}{p})$. H *acts with two orbits on* $\widehat{N}$, *namely* $\{\phi_0\}$ *and* $\{\phi_k\}_{k=1}^{p-1}$. *Then* $K_0 = Stab_H(\phi_0) = H$ *and* $K_1 = Stab_H(\phi_1) = \{1\}$. *The characters of* $K_0 = H = \mathbf{F}_p^*$ *are given as follows. Let* $\xi = \exp(\frac{2\pi i}{p-1})$ *and let* θ *be a multiplicative generator of* $\mathbf{F}_p^*$. *For* $0 \leq m \leq p-2$ *define* $\psi_m(\theta^t) = \xi^{mt}$. *Then* $\widehat{H} = \{\psi_m\}_{m=0}^{p-2}$. *The resulting induced representations* $\eta_m = \mathrm{Ind}_{K_0 \ltimes N}^G (\psi_m \ltimes \phi_0) = \psi_m \ltimes \phi_0$ *are given by*

$$\eta_m \left(\begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} \right) = \psi_m(a).$$

We next compute $\rho = \mathrm{Ind}_{K_1 \ltimes N}^G (\mathbf{1} \ltimes \phi_1) = \mathrm{Ind}_N^G \phi_1$. *For* $h \in H$ *write* $\tilde{h} = \begin{bmatrix} h & 0 \\ 0 & 1 \end{bmatrix} \in G$. *The elements of* $H = \{\tilde{h} : h \in H\}$ *are cosets representatives for* N *in* G, *and* $\{e_{\tilde{h},1} : h \in H\}$ *is a basis of* $C_{N,\phi_1}(G, \mathbf{C})$. *Let* $g = \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} \in G$ *and* $h \in H$. *Then*

$$\pi(g, \tilde{h}) = \pi \left(\begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} h & 0 \\ 0 & 1 \end{bmatrix} \right) = \begin{bmatrix} ah & 0 \\ 0 & 1 \end{bmatrix} = \widetilde{ah}$$

and

$$\pi(g, \tilde{h})^{-1} g \tilde{h} = \begin{bmatrix} 1 & (ah)^{-1}b \\ 0 & 1 \end{bmatrix}.$$

Hence

$$\rho(g) e_{\tilde{h},1} = \phi_1((ah)^{-1}b) e_{\widetilde{ah},1}.$$

2 Small Oscillations and Symmetry

In this section we describe an application of representation theory to the dynamics of mass-spring systems. In subsection 2.1 we recall the Euler-Lagrange equation for critical paths of the action functional. The Hamilton principle and its application to small oscillations are discussed in subsections 2.2 and 2.3. In subsection 2.4 we focus on mass-spring system, derive their equations of motion and study some examples. In subsection **??** we introduce vector bundles over finite spaces, and discuss some of their representation theoretic aspects. Finally, in subsection **??** we describe a representation theoretic method that substantially simplifies the determination of the normal modes of a mass-spring system.

2.1 The Euler-Lagrange Equation

Let $F : \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^n$ be a smooth function. Let $p_0, p_1 \in \mathbf{R}^n$ be fixed points, and let $t_0 < t_1 \in \mathbf{R}$ be fixed times. Consider the set of smooth paths

$$P_{t_0,p_0}^{t_1,p_1} = \{\gamma : [t_0, t_1] \to \mathbf{R}^n : \gamma(t_0) = p_0, \gamma(t_1) = p_1\}.$$

The *action functional* associated to F is the map $A : P_{t_0,p_0}^{t_1,p_1} \to \mathbf{R}$ given by

$$A(\gamma) = \int_{t=t_0}^{t_1} F(t, \gamma(t), \dot{\gamma}(t))dt.$$

Suppose γ is a minimum of A:

$$A(\gamma) = \min \left\{ A(\tilde{\gamma}) : \tilde{\gamma} \in P_{t_0,p_0}^{t_1,p_1} \right\}. \tag{15}$$

Let

$$h \in S_{t_0}^{t_1} = \{h : [t_0, t_1] \to \mathbf{R}^n : h(t_0) = h(t_2) = 0\},$$

and let $\lambda \in \mathbf{R}$. The path $\gamma + \lambda h$ belongs to $P_{t_0,p_0}^{t_1,p_1}$. define $g_h : (-1, 1) \to \mathbf{R}$ by $g_h(\lambda) = A(\gamma + \lambda h)$. the assumption (15) implies that $g_h(\lambda)$ has a minimum in $\lambda = 0$ and therefore $\frac{dg_h}{d\lambda}(0) = 0$. We say that $\gamma \in P_{t_0,p_0}^{t_1,p_1}$ is a *critical path* for the action A if $\frac{dg_h}{d\lambda}(0) = 0$ for all $h \in S_{t_0}^{t_1}$.

Proposition 2.1. *If $\gamma \in P_{t_0,p_0}^{t_1,p_1}$ is critical then it satisfies the Euler-Lagrange equation:*

$$\frac{\partial F}{\partial x}(t, \gamma(t), \dot{\gamma}(t)) = \frac{d}{dt}\left(\frac{\partial F}{\partial v}(t, \gamma(t), \dot{\gamma}(t))\right). \tag{16}$$

Proof.

$$g_h(\lambda) = A(\gamma + \lambda h) = \int_{t=t_0}^{t_1} F\left(t, \gamma(t) + \lambda h(t), \dot{\gamma}(t) + \lambda \dot{h}(t)\right) dt. \tag{17}$$

Differentiating (17) by λ we obtain

$$\frac{dg_h}{d\lambda}(\lambda) =$$
$$\int_{t=t_0}^{t_1} \left[\frac{\partial F}{\partial x}\left(t, \gamma(t) + \lambda h(t), \dot{\gamma}(t) + \lambda \dot{h}(t)\right) \cdot h(t) + \frac{\partial F}{\partial v}\left(t, \gamma(t) + \lambda h(t), \dot{\gamma}(t) + \lambda \dot{h}(t)\right) \cdot \dot{h}(t)\right] dt.$$

Therefore

$$\begin{aligned} 0 &= \frac{dg_h}{d\lambda}(0) \\ &= \int_{t=t_0}^{t_1} \left[\frac{\partial F}{\partial x}(t, \gamma(t), \dot{\gamma}(t)) \cdot h(t) + \frac{\partial F}{\partial v}(t, \gamma(t), \dot{\gamma}(t)) \cdot \dot{h}(t)\right] dt. \end{aligned} \tag{18}$$

Evaluating the second terms on the right hand side of (18) using integration by parts, we obtain:

$$\begin{aligned} &\int_{t=t_0}^{t_1} \frac{\partial F}{\partial v}(t, \gamma(t), \dot{\gamma}(t)) \cdot \dot{h}(t)dt \\ &= \left[\frac{\partial F}{\partial v}(t, \gamma(t), \dot{\gamma}(t)) \cdot h(t)\right]_{t=t_0}^{t=t_1} - \int_{t=t_0}^{t_1} \frac{d}{dt}\left(\frac{\partial F}{\partial v}(t, \gamma(t), \dot{\gamma}(t))\right) \cdot h(t)dt \\ &= -\int_{t=t_0}^{t_1} \frac{d}{dt}\left(\frac{\partial F}{\partial v}(t, \gamma(t), \dot{\gamma}(t))\right) \cdot h(t)dt. \end{aligned} \tag{19}$$

Combining (18) and (19) we get

$$0 = \int_{t=t_0}^{t_1} \left[\frac{\partial F}{\partial x}(t, \gamma(t), \dot{\gamma}(t)) - \frac{d}{dt}\left(\frac{\partial F}{\partial v}(t, \gamma(t), \dot{\gamma}(t)) \right) \right] \cdot h(t) dt. \tag{20}$$

As (20) holds for all $h \in S_{t_0}^{t_1}$, it follows that

$$\frac{\partial F}{\partial x}(t, \gamma(t), \dot{\gamma}(t)) = \frac{d}{dt}\left(\frac{\partial F}{\partial v}(t, \gamma(t), \dot{\gamma}(t)) \right).$$

□

We next formulate an invariance property of the Euler-Lagrange equation. Suppose $f = (f_1, \ldots, f_n) : \mathbf{R}^n \to \mathbf{R}^n$ is a diffeomorphism of $\mathbf{R}^n$. Let $Df(x)$ be the differential of f, i.e.

$$Df(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}.$$

Given a function $G = G(u, v) : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$, define $F = F(x, v) : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$, by $F(x, v) = G(f(x), Df(x)v)$.

Claim 2.2. *Let $q : [t_0, t_1] \to \mathbf{R}^n$, and let $h : [t_0, t_1] \to \mathbf{R}^n$ be given by $h(t) = f(q(t))$. Then*

$$\frac{\partial F}{\partial x}(q(t), \dot{q}(t)) = \frac{d}{dt}\left(\frac{\partial F}{\partial v}(q(t), \dot{q}(t)) \right) \tag{21}$$

iff

$$\frac{\partial G}{\partial y}(h(t), \dot{h}(t)) = \frac{d}{dt}\left(\frac{\partial G}{\partial u}(h(t), \dot{h}(t)) \right). \tag{22}$$

Proof. For notational convenience, we'll only give a proof for $n = 1$. The general case is similar. Suppose then that $G = G(y, u) : \mathbf{R}^2 \to \mathbf{R}$, and let $F(x, v) = G(f(x), f'(x)v)$. Then

$$\frac{\partial F}{\partial x}(x, v) = \frac{\partial G}{\partial y}(f(x), f'(x)v) \cdot f'(x) + \frac{\partial G}{\partial u}(f(x), f'(x)v) \cdot f''(x)v$$

and

$$\frac{\partial F}{\partial v}(x, v) = \frac{\partial G}{\partial u}(f(x), f'(x)v) \cdot f'(x).$$

Substituting $x = q(t)$ and $v = \dot{q}(t)$ and noting that $f'(q(t))\dot{q}(t) = \dot{h}(t)$, it follows that

$$\begin{aligned} &\frac{\partial F}{\partial x}(q(t), \dot{q}(t)) \\ &= \frac{\partial G}{\partial y}(h(t), \dot{h}(t)) \cdot f'(q(t)) + \frac{\partial G}{\partial u}(h(t), \dot{h}(t)) \cdot f''(q(t))\dot{q}(t) \end{aligned} \tag{23}$$

and

$$\frac{\partial F}{\partial v}(q(t), \dot{q}(t)) = \frac{\partial G}{\partial u}(h(t), \dot{h}(t)) \cdot f'(q(t)). \tag{24}$$

Differentiating (24) we obtain

$$\begin{aligned} &\frac{d}{dt}\left(\frac{\partial F}{\partial v}(q(t), \dot{q}(t))\right) \\ &= \frac{d}{dt}\left(\frac{\partial G}{\partial u}(h(t), \dot{h}(t))\right) \cdot f'(q(t)) + \frac{\partial G}{\partial u}(h(t), \dot{h}(t)) \cdot f''(q(t))\dot{q}(t). \end{aligned} \tag{25}$$

Subtracting (25) from (23) we get

$$\begin{aligned} &\frac{\partial F}{\partial x}(q(t), \dot{q}(t)) - \frac{d}{dt}\left(\frac{\partial F}{\partial v}(q(t), \dot{q}(t))\right) \\ &= \left(\frac{\partial G}{\partial y}(h(t), \dot{h}(t)) - \frac{d}{dt}\left(\frac{\partial G}{\partial u}(h(t), \dot{h}(t))\right)\right) \cdot f'(q(t)) \end{aligned}$$

As $f'(q(t)) \neq 0$, this completes the proof of the Claim.

□

2.2 A Little Mechanics

Consider a mechanical system whose configuration space is an open subset $\Omega \subset \mathbf{R}^n$. For example, suppose we have N particles whose dynamics is determined by a certain force field. The location of each particle is specified by 3 coordinates, so the configuration space is (a subset of) $\mathbf{R}^{3N}$. Let $T(x, v)$ be the kinetic energy and let $V(t, x)$ be the potential energy of the system. The Lagrangian of the system is

$$L(t, x, v) = T(x, v) - V(t, x).$$

The *Hamilton Principle* asserts that if $q \in P_{t_0,p_0}^{t_1,p_1}$ is a time evolution of the system, then q is a critical path of the action

$$A(q) = \int_{t=t_0}^{t_1} L(t, q(t), \dot{q}(t))dt.$$

In particular, q satisfies the Euler-Lagrange equation

$$\frac{\partial L}{\partial x}(t, q(t), \dot{q}(t)) = \frac{d}{dt}\left(\frac{\partial L}{\partial v}(t, q(t), \dot{q}(t))\right). \tag{26}$$

Example 2.3 (Newton's Second Law)**.** *The second law $F = ma$ is a special case of Hamilton principle. Indeed, consider a particle of mass m that moves under a conservative force field $F = -\nabla V$, where $V(x) = V(x_1, x_2, x_3)$ is the potential function. The kinetic energy is $T(x, v) = \frac{1}{2}m|v|^2$, hence $L(x, v) = \frac{1}{2}m|v|^2 - V(x)$. Now*

$$\frac{\partial L}{\partial x}(x, v) = -\nabla V = F,$$

and

$$\frac{\partial L}{\partial v}(x, v) = mv.$$

It follows that

$$F(q(t)) = \frac{d}{dt}(m\dot{q}(t)) = m\ddot{q}(t).$$

Example 2.4 (The Harmonic Oscillator)**.**

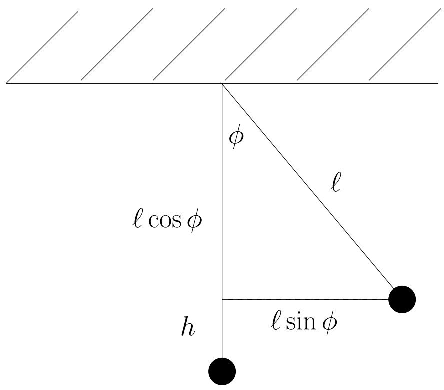


Figure 2: Simple pendulum

Example 2.5 (The simple pendulum)**.** *Here* $(x(t), y(t)) = (\ell \sin \phi(t), \ell \cos \phi(t))$. *Therefore* $|v(t)|^2 = \dot{x}(t)^2 + \dot{y}(t)^2 = \ell^2 \dot{\phi}(t)^2$. *Thus the kinetic energy is* $T(\phi) = \frac{1}{2} m\ell^2 \dot{\phi}^2$. *The potential energy is* $V(\phi) = mgh = mg\ell(1 - \cos \phi)$. *The Lagrangian is therefore*

$$L = T - V = \frac{1}{2} m\ell^2 \dot{\phi}^2 - mg\ell(1 - \cos \phi).$$

The Euler-Lagrange equation is:

$$- mg\ell \sin \phi = \frac{\partial L}{\partial \phi} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\phi}}\right) = m\ell^2 \ddot{\phi}. \tag{27}$$

Therefore

$$\ddot{\phi} = -\frac{g \sin \phi}{\ell}.$$

2.3 Small Oscillations - General Theory

Consider a mechanical system with configuration space $\mathbf{R}^n$. Let $T = \frac{1}{2}(M(x)v, v)$ denote the kinetic energy of the system, where $M(x)$ is a symmetric positive definite matrix that depends on the configuration x alone (and not on t). Let $V = P(x)$ denote the potential energy of the system. The Lagrangian of the system is

$$\mathcal{L}(x, v) = T - V = \frac{1}{2}(M(x)v, v) - P(x).$$

Suppose now that $q_0 \in \mathbf{R}^n$ is a stable equilibrium of the system. In particular:
(i) The force field at q_0 is zero, i.e.

$$0 = \nabla P(q_0) = \left(\frac{\partial P}{\partial x_1}(q_0), \ldots, \frac{\partial P}{\partial x_n}(q_0) \right). \tag{28}$$

(ii) The stability implies that the Hessian of P at q_0

$$K = \begin{bmatrix} \frac{\partial^2 P}{\partial x_1^2}(q_0) & \cdots & \frac{\partial^2 P}{\partial x_1 \partial x_n}(q_0) \\ \vdots & & \vdots \\ \frac{\partial^2 P}{\partial x_n \partial x_1}(q_0) & \cdots & \frac{\partial^2 P}{\partial x_n^2}(q_0) \end{bmatrix} \tag{29}$$

is positive semidefinite. Next note that

$$(M(q_0 + x)v, v) = (M(q_0)v, v) + O(|x| \cdot |v|^2) \tag{30}$$

and by the Taylor approximation and (28)

$$P(q_0 + x) = P(q_0) + \frac{1}{2}(Kx, x) + O(|x|^3). \tag{31}$$

Writing $M = M(q_0)$ it follows that

$$\mathcal{L}(q_0 + x, v) = \frac{1}{2}(Mv, v) - \frac{1}{2}(Kx, x) - P(q_0) + O(|x| \cdot (|x|^2 + |v|^2)).$$

For small $|x|, |v|$ we will therefore replace $\mathcal{L}(q_0 + x, v)$ by the linearized Lagrangian

$$L(x, v) = \frac{1}{2}(Mv, v) - \frac{1}{2}(Kx, x).$$

Claim 2.6. *A critical path $q(t)$ of the action functional $\int L(q(t), \dot{q}(t))dt$ satisfies*

$$M\ddot{q} = -Kq. \tag{32}$$

Proof. Note that $\frac{\partial L}{\partial x}(x, v) = -Kx$ and $\frac{\partial L}{\partial v}(x, v) = Mv$. It follows by the Euler-Lagrange equation that

$$\begin{aligned} -Kq &= \frac{\partial L}{\partial x}(q, \dot{q}) = \frac{d}{dt}\left(\frac{\partial L}{\partial v}(q, \dot{q}) \right) \\ &= \frac{d}{dt}(M\dot{q}) = M\ddot{q}. \end{aligned}$$

□

Let $H = M^{-\frac{1}{2}} K M^{-\frac{1}{2}}$. Then H is again symmetric positive semidefinite. Let $\omega_1^2, \ldots \omega_\ell^2$ be the eigenvalues of H, where ω_i appears with multiplicity m_i. Let $U_i = \{u \in \mathbf{R}^n : Hu = \omega_i^2 u\}$. Then $\mathbf{R}^n = U_1 \oplus \cdots \oplus U_t$ is an orthogonal decomposition. Note that if $u \in U_i$ then $M^{-\frac{1}{2}} K (M^{-\frac{1}{2}} u) = Hu = w_i^2 u$, and hence

$$K\left(M^{-\frac{1}{2}} u\right) = w_i^2 M^{\frac{1}{2}} u = w_i^2 M \left(M^{-\frac{1}{2}} u\right).$$

It follows that $V_i = M^{-\frac{1}{2}} U_i$ satisfies $V_i = \{v \in \mathbf{R}^n : Kv = w_i^2 Mv\}$ and that $\mathbf{R}^n = V_1 \oplus \cdots \oplus V_\ell$. The elements of V_i are called *normal modes* of the system, and they give rise to the basic solutions of (32).

Claim 2.7. *Let $\beta \in \mathbf{R}$ and let $v \in V_i$.*
(i) If $\omega_i \neq 0$ then $q(t) = \sin(w_i t + \beta) v$ satisfies (32).
(ii) If $\omega_i = 0$ then $q(t) = (t + \beta) v$ satisfies (32).
(iii) Any solution of (32) is a linear combination of the solutions given in (i) and (ii).

Proof. (i) If $\omega_i \neq 0$ then for $q(t) = \sin(w_i t + \beta) v$

$$M\ddot{q}(t) = -w_i^2 \sin(w_i t + \beta) Mv = -\sin(w_i t + \beta) Kv = -Kq(t). \tag{33}$$

(ii) If $\omega_i = 0$ then for $q(t) = (t + \beta) v$

$$M\ddot{q}(t) = 0 = -(t + \beta) Kv = -Kq(t). \tag{34}$$

(iii) Later.

□

Remark: Instead of finding the eigenvalues and eigenvectors of H, it is sometimes more convenient to find directly the ω_i^2 that satisfy $\det(K - \omega_i^2 M) = 0$ and then to compute $V_i = \ker(K - \omega_i^2 M)$.

2.4 Mass-Spring Systems

Consider N point masses $m_1, \ldots, m_N$ in $\mathbf{R}^d$, where m_i is located at $p_i \in \mathbf{R}^d$. Let $([N], E)$ be a graph that specifies which pairs are connected by a spring. For $\{\alpha, \beta\} \in E$, let $k_{\alpha\beta}$ be the corresponding spring constant. The kinetic energy of the system is

$$T(v) = \frac{1}{2} \sum_{i=1}^{N} m_i v_i^2 \tag{35}$$

and the potential energy is

$$P(x) = \frac{1}{2} \sum_{\{\alpha,\beta\} \in E} k_{\alpha\beta} \left(|(p_\alpha + x_\alpha) - (p_\beta + x_\beta)| - |p_\alpha - p_\beta|\right)^2. \tag{36}$$

We assume that $x = 0$ is a stable equilibrium of the system, i.e. with the masses in locations $p_1, \ldots, p_N$, the springs are relaxed. For $1 \leq \alpha \leq N$, let $p_\alpha = (p_{\alpha 1}, \ldots, p_{\alpha d})$. The $Nd \times Nd$ matrices M and K are given by the following

Proposition 2.8. *The matrix M is a block matrix $M = (M_{\alpha\beta})_{\alpha,\beta=1}^n$ where $M_{\alpha,\beta} = \delta_{\alpha,\beta} m_\alpha I_d$. The matrix K is a block matrix $K = (K_{\alpha\beta})_{\alpha,\beta=1}^n$ where $K_{\alpha\beta}$ is the $d \times d$ matrix given by*

$$K_{\alpha\beta} = \begin{cases} \sum\limits_{\gamma \in \Gamma(\alpha)} \dfrac{k_{\alpha\gamma}(p_\alpha - p_\gamma)(p_\alpha - p_\gamma)^t}{|p_\alpha - p_\gamma|^2} & \alpha = \beta, \\ -\frac{k_{\alpha\beta}(p_\alpha - p_\beta)(p_\alpha - p_\beta)^t}{|p_\alpha - p_\beta|^2} & \{\alpha, \beta\} \in E, \\ 0 & \textit{otherwise.} \end{cases} \tag{37}$$

Proof. The statement concerning M is clear. We proceed to compute $K = \mathrm{Hess}(P)(0)$. We first consider the case of two masses α, β in locations $p_\alpha \neq p_\beta$ with spring constant $k_{\alpha\beta}$. The potential function for this pair is given by

$$P_{\alpha\beta}(x_\alpha, x_\beta) = \frac{k_{\alpha\beta}}{2} \left(|\, (p_\alpha + x_\alpha) - (p_\beta + x_\beta) \,| - |p_\alpha - p_\beta| \right)^2 .$$

Note that for fixed $0 \neq u \in \mathbf{R}^n$

$$(|u + z| - |u|)^2 = \frac{(u \cdot z)^2}{u \cdot u} + O\left(|z|^3\right). \tag{38}$$

Applying (38) with $u = p_\alpha - p_\beta$ and $z = x_\alpha - x_\beta$ we obtain

$$P_{\alpha\beta}(x_\alpha, x_\beta) = \frac{k_{\alpha\beta}}{2} \cdot \frac{\left((p_\alpha - p_\beta) \cdot (x_\alpha - x_\beta)\right)^2}{|p_\alpha - p_\beta|^2}$$

$$= \frac{k_{\alpha\beta}[x_\alpha, x_\beta]}{2|p_\alpha - p_\beta|^2} \cdot \begin{bmatrix} (p_\alpha - p_\beta)(p_\alpha - p_\beta)^t & -(p_\alpha - p_\beta)(p_\alpha - p_\beta)^t \\ -(p_\alpha - p_\beta)(p_\alpha - p_\beta)^t & (p_\alpha - p_\beta)(p_\alpha - p_\beta)^t \end{bmatrix} \cdot \begin{bmatrix} x_\alpha \\ x_\beta \end{bmatrix}. \tag{39}$$

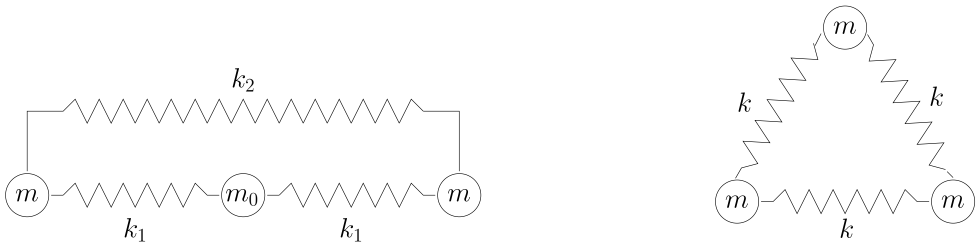
Using (39) for all pairs (α, β) it follows that

$$P(x) = \frac{1}{2} x^t K x + O(|x|^3)$$

and hence $\mathrm{Hess}(P)(0) = K$.

$\square$

Using Proposition 2.8 and Claim 2.7 one can, in principle, determine the motion of mass-spring systems. We give two small examples.



(a) System with $d = 1$ (b) System with $d = 2$

Figure 3

Example 2.9. *Consider the system of three collinear masses depicted in Figure 3(a). By Proposition 2.8, the matrices M and K are given by*

$$M = \begin{bmatrix} m & 0 & 0 \\ 0 & m_0 & 0 \\ 0 & 0 & m \end{bmatrix} \quad , \quad K = \begin{bmatrix} k_1 + k_2 & -k_1 & -k_2 \\ -k_1 & 2k_1 & -k_1 \\ -k_2 & -k_1 & k_1 + k_2 \end{bmatrix}.$$

To find the eigenvalues ω^2, we solve

$$\det(K - \omega^2 M) = \det \begin{bmatrix} k_1 + k_2 - m\omega^2 & -k_1 & -k_2 \\ -k_1 & 2k_1 - m_0\omega^2 & -k_1 \\ -k_2 & -k_1 & k_1 + k_2 - m\omega^2 \end{bmatrix} = 0.$$

The solutions are:

(i) $\omega_1^2 = 0$ *with normal mode* $v_1 = (1, 1, 1)$. *The corresponding solution of (32) is* $q_1(t) = (t + \beta)v_1$, *i.e. the system moves uniformly.*

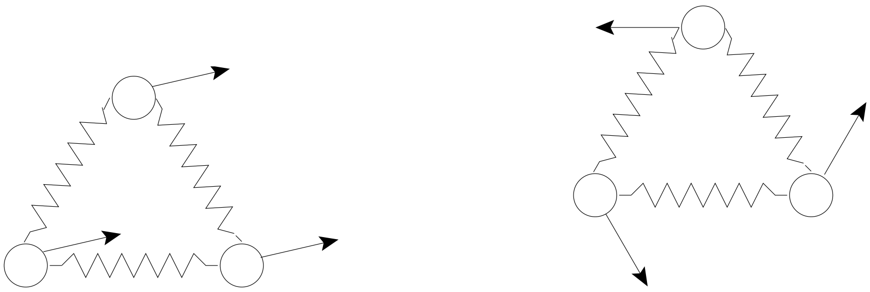
(ii) $\omega_2^2 = \frac{2k_1}{m_0} + \frac{k_1}{m}$ *with normal mode* $v_2 = (1, -\frac{2m}{m_0}, 1)$, *and* $q_2(t) = \sin(\omega_2 t + \beta)v_2$. *Thus the two m's move in one direction, and m_0 moves in the other direction.*

(iii) $\omega_3^2 = \frac{k_1 + 2k_2}{m}$ *with normal mode* $v_3 = (1, 0, -1)$, *and* $q_3(t) = \sin(\omega_3 t + \beta)v_3$. *Here the two m's move in opposite directions, while m_0 is stationary.*

Example 2.10. *Consider the system of three masses located at the vertices of an equilateral triangle depicted in Figure 3(b). By Proposition 2.8, the matrices M and K are given by $M = mI_6$ and*

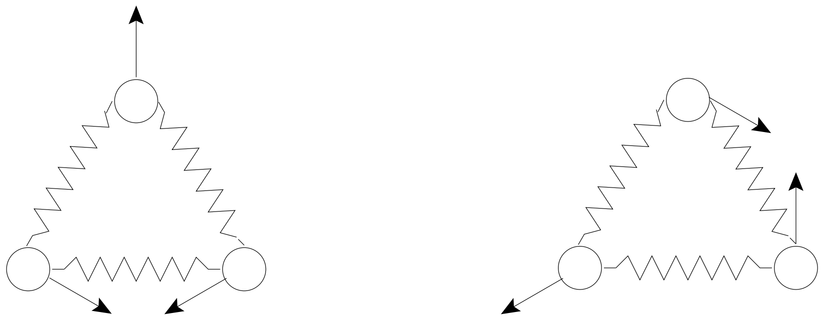
$$K = \frac{k}{4} \begin{bmatrix} 5 & \sqrt{3} & -4 & 0 & -1 & -\sqrt{3} \\ \sqrt{3} & 3 & 0 & 0 & -\sqrt{3} & -3 \\ -4 & 0 & 5 & -\sqrt{3} & -1 & \sqrt{3} \\ 0 & 0 & -\sqrt{3} & 3 & \sqrt{3} & -3 \\ -1 & -\sqrt{3} & -1 & \sqrt{3} & 2 & 0 \\ -\sqrt{3} & -3 & \sqrt{3} & -3 & 0 & 6 \end{bmatrix}.$$

To find the ω_i^2 we solve $\det(K - \omega^2 M) = 0$. *The solutions are:*



(a) Translation mode (b) Rotation mode

Figure 4: $w_1^2 = 0$



(a) (b)

Figure 5: $\omega_2^2 = \frac{3k}{2m}$ - vibratory mode

(i) $\omega_1^2 = 0$*. Then* $V_1 = \{v \in \mathbf{R}^6 : Kv = 0\}$ *is a* 3*-dimensional spanned by the vectors*

$$u_1 = (1, 0, 1, 0, 1, 0)\ ,\ u_2 = (0, 1, 0, 1, 0, 1)\ ,\ u_3 = (-1, \sqrt{3}, -1, -\sqrt{3}, 2, 0).$$

The corresponding solutions of (32), namely $q_1(t) = (t + \beta)v$ *where* $v \in V_1$*, describe two kinds of uniform motions. If* $v \in V_{1,T} = span\{u_1, u_2\}$ *then all masses move in the same direction in the plane - see Figure 4(a). For* $v \in V_{1,R} = span\{u_3\}$*, the masses rotate around the center of the triangle - see Figure 4(b).*

(ii) $\omega_2^2 = \frac{3k}{2m}$*. Then* $V_2 = \{v \in \mathbf{R}^6 : (K - \frac{3k}{2m}M)v = 0\}$ *is a* 2*-dimensional space spanned by the vectors*

$$u_1 = (\sqrt{3}, -1, -\sqrt{3}, -1, 0, 2)\ ,\ u_2 = (-\sqrt{3}, -1, 0, 2, \sqrt{3}, -1).$$

The vibratory modes corresponding to u_1, u_2 *are depicted in Figure 5 (a) and (b).*

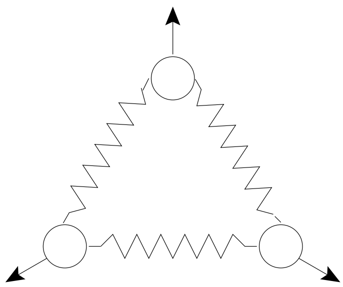


Figure 6: $\omega_3^2 = \frac{3k}{m}$ - vibratory mode

(iii) $\omega_3^2 = \frac{3k}{m}$. *Then* $V_3 = \{v \in \mathbf{R}^6 : (K - \frac{3k}{m}M)v = 0\}$ *is a* 1*-dimensional space spanned by* $u = (-\sqrt{3}, -1, \sqrt{3}, -1, 0, 2)$. *The corresponding vibratory mode is depicted in Figure 6.*

2.5 Mass-Spring Systems with Symmetry

The computation of normal modes using eigenvalues as above can sometimes be simplified considerably using representation theory. In Example 1.16 we associated with a finite $X \subset \mathbf{R}^d$ a discrete vector bundle $E = \cup_{x \in X} E_x$, where $E_x = \{x\} \times \mathbf{C}^d$, with the natural projection map $E_x \to x$. Suppose now that $G \subset \mathrm{Aut}(X) = \{g \in O(d) : gX = X\}$. The action of G on E defined by $g(x, v) = (gx, gv)$, gives rise to a representation ρ on $\Gamma(E)$ given by

$$\rho(g)(s)(x) = gs(g^{-1}x).$$

For $g \in G$ let $\mathrm{Fix}(g) = \{x \in X : gx = x\}$. Then by Corollary 1.19, the character χ_ρ of ρ satisfies

$$\chi_\rho(g) = \sum_{\{x : gx = x\}} \mathrm{tr}[g : Ex \to E_x] = |\mathrm{Fix}(g)| \cdot \mathrm{tr}(g). \tag{40}$$

The subspace $\Gamma_T(E) \subset \Gamma(E)$ of translation sections is defined by

$$\Gamma_T(E) = \{s \in \Gamma(E) : s(x) = s(y) \text{ for all } x, y \in X\}.$$

Clearly, $\Gamma_T(E)$ is a d-dimensional G-invariant subspace of $\Gamma(E)$. Denote by ρ_T the restriction of ρ to $\Gamma_T(E)$. Then

$$\chi_{\rho_T}(g) = \mathrm{tr}(g). \tag{41}$$

The subspace $\Gamma_R(E) \subset \Gamma(E)$ of rotation sections is defined by

$$\Gamma_R(X) = \{s \in \Gamma(E) : s \text{ linear }, x \cdot sx = 0 \text{ for all } x \in X\}.$$

Clearly, $\Gamma_R(X)$ is a G-invariant subspace of $\Gamma(E)$.

Consider a spring-mass system in $\mathbf{R}^d$, i.e. a set $X = \{p_\alpha\}_{\alpha=1}^n \subset \mathbf{R}^d$ with masses $m(p_\alpha) = m_\alpha$ and spring constants $k(p_\alpha, p_\beta) = k_{\alpha\beta}$ for $1 \leq \alpha, \beta \leq n$. The small oscillations dynamics of the system is driven by the Euler-Lagrange equation

$$M\ddot{q}(t) = -Kq(t), \tag{42}$$

where M, K are given in Proposition 2.8 A *symmetry of the system* is a $T \in O(d)$ such that $TX = X$, $m(Tp_\alpha) = m_\alpha$ and $k(Tp_\alpha, Tp_\beta) = k_{\alpha\beta}$. Let G denote the group of symmetries of the system. Recall that $V_i = \ker(K - \omega_i^2 M)$ is the space of normal modes corresponding to ω_i^2. Let $\Gamma(E) = \oplus_{i=1}^{\ell} V_i$ the direct sum decomposition of $\Gamma(E)$ into normal modes subspaces.

Claim 2.11. *V_i is invariant under G.*

Proof. Let $T \in G$ and $v \in V_i$. If $\omega_i \neq 0$ then $q(t) = \sin(\omega_i t)v$ satisfies (42). Clearly $h(t) = Tq(t)$ also satisfies (42), and therefore

$$\begin{aligned} -\sin(\omega_i t)\omega_i^2 MTv &= M\left(-\sin(\omega_i t)\omega_i^2 Tv\right) \\ &= M\ddot{h}(t) = -Kh(t) = -\sin(\omega_i t)KTv. \end{aligned}$$

It follows that $K(Tv) = \omega_i^2 M(Tv)$ and hence $Tv \in V_i$. The case $\omega_i = 0$ is similar.

□

Let $\{W_j\}_{j=1}^t$ be the irreducible representations of G. Let $\Gamma(E) = \oplus_{j=1}^t L_j$ where L_j is the isotypic component of $\Gamma(E)$ corresponding to W_j.

Corollary 2.12. *If L_j is irreducible, then $L_j \subset V_i$ for some i.*

This suggests the following approach to determining at least some of the normal modes:

- Determine χ_ρ using (40).
- Decompose $\chi_\rho = \sum_{i=1}^t m_i \chi_i$ where $\{\chi_i\}_{i=1}^t$ are irreducible characters of G, and $m_i > 0$.
- Let $\Gamma(E) = \oplus_{j=1}^t W_j$ where W_j is the isotypic component of $\Gamma(E)$ corresponding to the character χ_j. Determine W_j by using the projection $P_j : \Gamma(E) \to W_j$ given by
$$P_j s = \frac{\chi_j(1)}{|G|} \sum_{g \in G} \overline{\chi_j(g)} \rho(g) s. \tag{43}$$
- If $m_j = 1$, then W_j is contained in some normal mode V_i. We can then recover ω_i^2 and the full V_i. Otherwise, a subspace of W_j is subspace of a normal mode and finding it may require additional considerations.

We now revisit Example 2.10 using symmetry.

Example 2.13. *The symmetry group of the equilateral triangle is $G = S_3$. The group of rotational symmetries is $G_R = C_3$. Let ρ_1, ρ_2, ρ_3 denote respectively the trivial, sign and standard representations, and let $\chi_i = \chi_{\rho_i}$. The character table of S_3 is*

	1	(12)	(123)
χ_1	1	1	1
χ_2	1	-1	1
χ_3	2	0	-1

We next compute χ_ρ and χ_{ρ_T} using (40) and (41).

	1	(12)	(123)
χ_ρ	6	0	0
χ_{ρ_T}	2	0	-1

It follows that $\chi_\rho = \chi_1+\chi_2+2\chi_3$ and $\chi_{\rho_T} = \chi_3$. Using the notations of Example 2.10, we have that $\Gamma_T(E) = V_{1,T}$ is the space of translational normal modes. Next consider the rotational normal modes. Clearly $\Gamma_R(E)$ is the 1-dimensional space generated by the rotation $R_{\frac{\pi}{2}}$. As $\rho(r)R_{\frac{\pi}{2}} = R_{\frac{\pi}{2}}$ and $\rho(s)R_{\frac{\pi}{2}} = -R_{\frac{\pi}{2}}$, it follows that $\Gamma_R(E)$ is isomorphic to the sign representation. Therefore the sum of the vibratory components in the decomposition of $\Gamma(E)$ has character

$$\chi_\rho - \chi_{\rho_T} - \chi_{\rho_R} = (\chi_1 + \chi_2 + 2\chi_3) - \chi_3 - \chi_2 = \chi_1 + \chi_3.$$

Let v_1, v_2, v_3 be the three vertices of an equilateral triangle with center 0, i.e. $v_1+v_2+v_3 = 0$. We first determine the normal mode corresponding to χ_1. Let $t = (12)$ be the reflection that switches v_1 and v_2. By (43), the normal mode corresponding to χ_1 is the image of $\Gamma(E)$ under the linear transformation

$$A = \sum_{g\in S_3} \rho(g) = \rho(1) + \rho(r) + \rho(r^2) + \rho(t) + \rho(tr) + \rho(tr^2).$$

Let $s \in \Gamma(E)$, and write $(s(v_1), s(v_2), s(v_3)) = (u_1, u_2, u_3)$. Then

$$\begin{aligned}(As(v_1), As(v_2), As(v_3)) = (u_1, u_2, u_3) + (ru_3, ru_1, ru_2) + (r^2u_2, r^2u_3, r^2u_1)\\ + (tu_2, tu_1, tu_3) + (tru_1, tru_3, tru_2) + (tr^2u_3, tr^2u_2, tr^2u_1).\end{aligned}$$

In particular, for $(s(v_1), s(v_2), s(v_3)) = (v_1, 0, 0)$ we obtain

$$\begin{aligned}(As(v_1), As(v_2), As(v_3)) &= (v_1, 0, 0) + (0, rv_1, 0) + (0, 0, r^2v_1)\\ &+ (0, tv_1, 0) + (trv_1, 0, 0) + (0, 0, tr^2v_1)\\ &= \big((I + tr)v_1, (r + t)v_1, (r^2 + tr^2)v_1\big) = 2(v_1, v_2, v_3).\end{aligned}$$

Thus the normal mode for χ_1 is spanned by the section $s \in \Gamma(E)$ given by

$$(s(v_1), s(v_2), s(v_3)) = (v_1, v_2, v_3).$$

This is of course the space V_3 *we computed in Example 2.10. Finally, we find the vibrational normal mode* V_2 *corresponding to* χ_3*. Let* W *be the* 4*-dimensional isotypic component of* $\Gamma(E)$ *corresponding to* χ_3*. We already know that one copy of* χ_3 *is* $\Gamma_T(E) = V_{1,T}$*. Therefore* $W = V_2 \oplus V_{1,T}$ *and (why?)* $V_2 \perp V_{1,T}$*. By (43),* $W \subset \Gamma(E)$ *given by the image of* $\Gamma(E)$ *under the linear transformation*

$$B = \sum_{g \in G} \overline{\chi_3(g)} \rho(g) = 2I - \rho(r) - \rho(r^2).$$

One can then compute W *and then* $V_2 = V_{1,T}^{\perp} \cap W$*. In more detail, let* $s \in \Gamma(E)$*, and write* $(s(v_1), s(v_2), s(v_3)) = (u_1, u_2, u_3)$*. Then*

$$\begin{aligned} &(Bs(v_1), Bs(v_2), Bs(v_3)) = 2(u_1, u_2, u_3) - (ru_3, ru_1, ru_2) - (r^2 u_2, r^2 u_3, r^2 u_1) \\ &= (2u_1 - ru_3 - r^2 u_2, 2u_2 - ru_1 - r^2 u_3, 2u_3 - ru_2 - r^2 u_1). \end{aligned}$$

In particular, for s *given by* $(s(v_1), s(v_2), s(v_3)) = (2v_1 + v_2, -(2v_1 + v_2), 0)$ *we obtain*

$$(Bs(v_1), Bs(v_2), Bs(v_3)) = 3(v_1, v_3, v_2)$$

and for s' *given by* $(s'(v_1), s'(v_2), s'(v_3)) = (-v_1 + v_2, v_1 - v_2, 0))$ *we obtain*

$$(Bs'(v_1), Bs'(v_2), Bs'(v_3)) = 3(v_2, v_1, v_3).$$

Then $V_2 = span\{s, s'\}$*, which of course coincides with the computation in Example 2.10. Note that we have determined the normal modes with no eigenvalue computation. Finally, to obtain the vibration frequencies* ω *corresponding to normal mode* s*, we simply compute* $Hs = \omega^2 s$ *where* $H = M^{-\frac{1}{2}} K M^{-\frac{1}{2}}$*.*

3 Quantum Systems

A closed *quantum system* is represented by a complex Hilbert space $\mathcal{H}$. We will assume, for the purposes of this section, that $\mathcal{H}$ is finite dimensional. At some points we will however comment on the physically more common case where $\mathcal{H}$ is infinite dimensional, e.g. $\mathcal{H} = L^2(\mathbf{R}^3)$. A *state* of the system is a 1-dimensional linear subspace of $\mathcal{H}$, i.e. $[\psi] = \text{span}\{\psi\}$ where $0 \neq \psi \in \mathcal{H}$. A *superposition* of states $[\psi_1], \ldots, [\psi_k]$ is a state of the form $[\sum_{i=1}^k a_i\psi_i]$. It is assumed that there is a bijective correspondence between physical quantity (energy, coordinates, momentum, etc.) and the space of self-adjoint operators on $\mathcal{H}$, i.e. linear operators $f : \mathcal{H} \to \mathcal{H}$ such that $f = f^*$. We thus identify such f as an *observable*. Fix a observable f on $\mathcal{H}$. Let $\lambda_1, \ldots, \lambda_t$ be the distinct eigenvalues of f with eigenspaces $\mathcal{H}_f(\lambda_1), \ldots, \mathcal{H}_f(\lambda_t)$. Note that the λ_i's are real and the $\mathcal{H}_f(\lambda_i)$'s are orthogonal. Let $p_i : \mathcal{H} \to \mathcal{H}_f(\lambda_i)$ denote the orthogonal projection of $\mathcal{H}$ onto $\mathcal{H}_f(\lambda_i)$. Let $\psi \in \mathcal{H}$ such that $|\psi| = 1$. A *measurement* of $[\psi]$ for the physical property that corresponds to f is a (non-deterministic) procedure that with probability $|p_k\psi|^2 = (\psi, p_k\psi)$ will

(a) Give the value λ_k for the measurement.

(b) Change the state $[\psi]$ into $[p_k\psi]$ immediately after the measurement.

Remarks:

- Note that $\sum_{k=1}^t p_k = I$ and $\sum_{k=1}^t \lambda_k p_k = f$. In particular
$$\sum_{k=1}^t |p_k\psi|^2 = \sum_{k=1}^t (\psi, p_k\psi) = (\psi, \psi) = 1.$$

- If f has $n = \dim \mathcal{H}$ distinct eigenvalues $\lambda_1, \ldots, \lambda_n$ with unit length eigenvectors $\psi_1, \ldots, \psi_n$, then $\mathcal{H}_f(\lambda_k) = \text{span}\{\psi_k\}$ and a measurement of the observable f will output the value λ_k and the state $[\psi_k]$ with probability $|p_k\psi|^2 = (\psi, p_k\psi)$. More generally, for $B \subset \mathbf{R}$ let $P_B = \sum_{\{k:\lambda_k \in B\}} p_k$. Then the probability that the output of the measurement of f on ψ will be in B is
$$(\psi, P_B\psi). \tag{44}$$

- If f, g are self-adjoint, then so are f^2, $f - \alpha$ for any $\alpha \in \mathbf{R}$ and $-i[f, g] = -i(fg - gf)$.

The *expectation* of an observable f with respect to the normalized state ψ (i.e. $|\psi| = 1$) is defined by
$$\langle f\rangle_\psi = \sum_k \lambda_k(\psi, p_k\psi) = \left(\psi, \left(\sum_k \lambda_k p_k\right)\psi\right) = (\psi, f\psi).$$

The *dispersion* of f in ψ is

$$\begin{aligned}\Delta f_\psi &= \sqrt{\langle (f-\langle f\rangle_\psi)^2\rangle_\psi} = \sqrt{\left((f-\langle f\rangle_\psi)^2\,\psi,\psi\right)} \\ &= \sqrt{\left((f-\langle f\rangle_\psi)\,\psi,(f-\langle f\rangle_\psi)\,\psi\right)} = |(f-\langle f\rangle_\psi)\,\psi| \\ &= \sqrt{|f\psi|^2-\langle f\rangle_\psi^2}.\end{aligned} \tag{45}$$

The *Lie bracket* of two linear operators f, g on $\mathcal{H}$ is $[f,g] = fg - gf$.

Proposition 3.1 (Heisenberg Uncertainty Principle)**.** *For any observables f, g*

$$\Delta f_\psi \cdot \Delta g_\psi \geq \frac{1}{2}\left|([f,g]\psi,\psi)\right|. \tag{46}$$

Proof: Let $f_1 = f - \langle f\rangle_\psi$ and $g_1 = g - \langle g\rangle_\psi$. Clearly, $[f_1, g_1] = [f, g]$. It follows that

$$\begin{aligned}|([f,g]\psi,\psi)| &= |([f_1,g_1]\psi,\psi)| = |(f_1g_1\psi,\psi)-(g_1f_1\psi,\psi)| \\ &\leq |(f_1g_1\psi,\psi)|+|(g_1f_1\psi,\psi)| \\ &= |(g_1\psi,f_1\psi)|+|(f_1\psi,g_1\psi)| \\ &\leq 2|f_1\psi|\cdot|g_1\psi| = 2\Delta f_\psi\cdot\Delta g_\psi.\end{aligned} \tag{47}$$

□

Example 3.2 (Particle on a Line)**.** *Here the relevant Hilbert space $\mathcal{H}$ is $L^2(\mathbf{R})$. Analysing some aspects of this system requires the use of the theory of distributions which is beyond the scope of this course. The arguments below should therefore be regarded as merely suggestive and not rigorous. The* position *observable is the operator $\widehat{x}$ given by $\widehat{x}\psi(x) = x\psi(x)$. One may view any x_0 as an eigenvalue of $\widehat{x}$ with (generalized) eigenfunction $\delta(x - x_0)$ (which may be thought of as the wave function of particle located at x_0). It follows that if $A \subset \mathbf{R}$ then the projection $P_{\widehat{x},A}$ of $\mathcal{H}$ on the subspace of wave functions with position supported in A satisfies $P_{\widehat{x},A}\psi = 1_A\psi$. Thus, if $|\phi| = 1$ then on measuring $\widehat{x}$ on ψ, the probability of obtaining a value in A is*

$$(\psi, P_{\widehat{x},A}\psi) = \int_{\mathbf{R}} \psi(x)\overline{P_{\widehat{x},A}\psi(x)}dx = \int_{\mathbf{R}} 1_A(x)|\psi(x)|^2dx = \int_{x\in A}|\psi(x)|^2dx.$$

The function $|\psi(x)|^2$ is thus the probability density of the position of the particle. The expectation of $\widehat{x}$ in the normalized state ψ is

$$\langle\widehat{x}\rangle_\psi = (\psi,\widehat{x}\psi) = \int_{x\in\mathbf{R}} x|\psi(x)|^2dx.$$

By (45), the dispersion of $\widehat{x}$ with respect to ψ is

$$\begin{aligned}\Delta\widehat{x}_\psi &= |\,(\widehat{x} - \langle\widehat{x}\rangle_\psi)\,\psi| \\ &= \sqrt{(\widehat{x}\psi - \langle\widehat{x}\rangle_\psi\psi, \widehat{x}\psi - \langle\widehat{x}\rangle_\psi\psi)} \\ &= \sqrt{|\widehat{x}\psi|^2 - \langle\widehat{x}\rangle_\psi^2} \\ &= \sqrt{\int_{x\in\mathbf{R}} |x|^2|\psi(x)|^2 dx - \left(\int_{x\in\mathbf{R}} x|\psi(x)|^2\right)^2}.\end{aligned} \tag{48}$$

In particular

$$\Delta\widehat{x}_\psi \leq \sqrt{\int_{x\in\mathbf{R}} |x|^2|\psi(x)|^2 dx}. \tag{49}$$

The momentum *operator $\widehat{p}$ is given by $\widehat{p}\psi(x) = -i\hbar\psi'(x)$. The eigenvalues of $\widehat{p}$ are $p \in \mathbf{R}$ with eigenfunctions $\psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} \exp\left(\frac{ipx}{\hbar}\right)$. Thus ψ_p is the wave function of a particle with constant momentum p. Let $\tilde{f}$ denote the Fourier transform of a function $f \in L^2(\mathbf{R})$, i.e.*

$$\widetilde{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{x\in\mathbf{R}} f(x)\exp(-ix\xi)dx.$$

The projection $P_{\widehat{p},B}$ of $\mathcal{H}$ on the space of wave functions with momentum supported in $B \subset \mathbf{R}$ satisfies

$$\begin{aligned}P_{\widehat{p},B}\psi &= \int_{p\in B} (\psi, \psi_p)\psi_p(x)dp \\ &= \frac{1}{\sqrt{2\pi\hbar}} \int_{p\in B} \left(\frac{1}{\sqrt{2\pi\hbar}} \int_{t\in\mathbf{R}} \psi(t)\exp\left(-\frac{ipt}{\hbar}\right) dt\right) \exp\left(\frac{ipx}{\hbar}\right) dp \\ &= \frac{1}{\sqrt{2\pi}\hbar} \int_{p\in B} \widetilde{\psi}\left(\frac{p}{\hbar}\right) \exp\left(\frac{ipx}{\hbar}\right) dp.\end{aligned}$$

Thus, if $|\psi| = 1$ then on measuring $\widehat{p}$ on ψ, the probability of obtaining a value in B is

$$\begin{aligned}(\psi, P_{\widehat{p},B}\psi) &= \int_{\mathbf{R}} \psi(x)\overline{P_{\widehat{p},B}\psi(x)}dx \\ &= \frac{1}{\hbar} \int_{x\in\mathbf{R}} \psi(x) \left(\frac{1}{\sqrt{2\pi}} \int_{p\in B} \overline{\widetilde{\psi}\left(\frac{p}{\hbar}\right)} \exp\left(-\frac{ipx}{\hbar}\right) dp\right) dx \\ &= \frac{1}{\hbar} \int_{p\in B} \left|\widetilde{\psi}\left(\frac{p}{\hbar}\right)\right|^2 dp.\end{aligned}$$

It follows that

$$\langle\widehat{p}\rangle_\psi = \frac{1}{\hbar} \int_{p\in\mathbf{R}} p \left|\widetilde{\psi}\left(\frac{p}{\hbar}\right)\right|^2 dp$$

and

$$\Delta\widehat{p}_{\psi} = \sqrt{\frac{1}{\hbar}\int_{p\in\mathbf{R}} p^2 \left|\widetilde{\psi}\left(\frac{p}{\hbar}\right)\right|^2 dp - \left(\frac{1}{\hbar}\int_{p\in\mathbf{R}} p \left|\widetilde{\psi}\left(\frac{p}{\hbar}\right)\right|^2 dp\right)^2}. \tag{50}$$

In particular

$$\begin{aligned}\Delta\widehat{p}_{\psi} &\le \sqrt{\frac{1}{\hbar}\int_{p\in\mathbf{R}} p^2 \left|\widetilde{\psi}\left(\frac{p}{\hbar}\right)\right|^2 dp} \\ &= \hbar\sqrt{\int_{p\in\mathbf{R}} p^2 \left|\widetilde{\psi}(p)\right|^2 dp.}\end{aligned} \tag{51}$$

It can be checked that $[\widehat{x}, \widehat{p}] = i\hbar I$. *Hence, if* $|\psi| = 1$ *then*

$$\Delta\widehat{x}_{\psi} \cdot \Delta\widehat{p}_{\psi} \ge \frac{\hbar}{2}.$$

Using (49) and (51) we obtain the Fourier theoretic version of the uncertainty inequality:

$$\|x\psi\| \cdot \|p\widetilde{\psi}\| \ge \frac{1}{2}. \tag{52}$$

The time evolution of a quantum system $\mathcal{H}$ depends on the *energy observable* or *Hamiltonian* of the system H, and can be viewed in two equivalent ways. Let ψ_0 be a state such that $|\psi_0| = 1$ and let f_0 be an observable. In the *Schrödinger picture*, f_0 remains constant, while ψ_0 develops in time, and the resulting curve of states $\{\psi_t\}_{t\in\mathbf{R}}$ satisfies the Schrödinger equation

$$i\hbar\frac{d\psi_t}{dt} = H\psi_t. \tag{53}$$

In other words,

$$\psi_t = \exp\left(-\frac{i}{\hbar}Ht\right)\psi_0. \tag{54}$$

Remark: Note that if $\alpha \in \mathbf{R}$ and H is self-adjoint, then $\exp(i\alpha H)$ is unitary. Indeed,

$$\begin{aligned}\exp(i\alpha H) \cdot (\exp(i\alpha H))^* &= \exp(i\alpha H) \cdot \exp(-i\alpha H^*) \\ &= \exp(i\alpha H) \cdot \exp(-i\alpha H) = I.\end{aligned}$$

In particular, $\exp(-\frac{i}{\hbar}Ht)$ is unitary.

We next describe the *Heisenberg picture* of quantum evolution. Here the state ψ_0 does not change, but the observable f_0 does and the resulting curve of observables f_t should satisfy

$$-i\hbar\frac{df_t}{dt} = [H, f_t]. \tag{55}$$

We next show that the two views of quantum evolutions are essentially equivalent.

Proposition 3.3. *Let $U(t)$ be a curve of unitary operators on $\mathcal{H}$. Let ψ_0 be a state such that $|\psi_0| = 1$ and let f_0 be an observable. Let $\psi_t = U(t)\psi_0$ and let $f_t = U(t)^{-1}f_0U(t)$. Then:*

(i) The expectation of f_0 in the state ψ_t is equal to the expectation of f_t in the state ψ_0:

$$\langle f_0\rangle_{\psi_t} = \langle f_t\rangle_{\psi_0}.$$

(ii) The following conditions are equivalent:

(a) For any initial state ψ_0, the curve $\psi_t = U(t)\psi_0$ satisfies (53).

(b) For any initial observable f_0, the curve $f_t = U(t)^{-1}f_0U(t)$ satisfies (55).

Proof. For (i) note that

$$\begin{aligned}\langle f_0\rangle_{\psi_t} &= (\psi_t, f_0\psi_t) = (U(t)\psi_0, f_0U(t)\psi_0)\\ &= (\psi_0, U(t)^{-1}f_0U(t)\psi_0) = (\psi_0, f_t\psi_0) = \langle f_t\rangle_{\psi_0}.\end{aligned} \tag{56}$$

Proof of (ii). Condition (a) is clearly equivalent to $i\hbar\dot{U}(t) = HU(t)$. On the other hand, note that

$$\frac{df_t}{dt} = [f_t, U(t)^{-1}\dot{U}(t)]. \tag{57}$$

Condition (b) is therefore equivalent to $-i\hbar[f_t, U(t)^{-1}\dot{U}(t)] = [H, f_t]$ for all initial observables f_0. Hence $i\hbar U(t)^{-1}\dot{U}(t) = H$, and again $i\hbar\dot{U}(t) = HU(t)$.

□

Example 3.4 (The Quantum Harmonic Oscillator)**.** *In Example 2.4 we discussed the classical harmonic oscillator whose Hamiltonian is $H_c(x,p) = \frac{p^2}{2m} + \frac{kx^2}{2}$. The general solution of the Hamilton equations is $x(t) = A\cos(\omega t + \alpha)$, where $\omega = \sqrt{\frac{k}{m}}$. The corresponding quantum system is $\mathcal{H} = L^2(\mathbf{R})$ with the quantized Hamiltonian*

$$H = H(\widehat{x}, \widehat{p}) = \frac{\widehat{p}^2}{2m} + \frac{k\widehat{x}^2}{2}$$

whose action of $\psi \in \mathcal{H}$ is given by

$$H\psi = -\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \frac{kx^2}{2}\psi(x).$$

It can be shown that the spectrum of H is $\lambda_n = \hbar\omega(n + \frac{1}{2})$ for integers $n \geq 0$.

Consider now n disjoint quantum systems $\mathcal{H}_1, \ldots, \mathcal{H}_n$. Our classical intuition may lead us to think that the Hilbert space $\mathcal{H}$ corresponding to the unified system is $\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n$. It turns out that in fact $\mathcal{H}$ is a subspace of $\bigotimes_{i=1}^n \mathcal{H}_i = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$. For simplicity, in the sequel we'll assume that $\mathcal{H} = \bigotimes_{i=1}^n \mathcal{H}_i$. The unified system $\mathcal{H}$ contains *decomposable states*, i.e. $\psi = \psi_1 \otimes \cdots \otimes \psi_n$, with $\psi_i \in \mathcal{H}_i$, that correspond to our intuition. However, the vast majority of states in $\mathcal{H}$ are *entangled*, i.e. not decomposable. Manipulation of such states is a key ingredient in quantum computation. The existence (indeed prevalence) of entangled states is a source of a number of highly non-intuitive phenomena. The states of the Hilbert space $\mathbf{C}^2$ are called *qubits*.

Example 3.5 (The Einstein-Podolosky-Rosen Paradox)**.** *Let $\mathcal{H}_1 = \mathcal{H}_2 = \mathbf{C}^2$ with the standard basis $e_1 = (1,0), e_2 = (0,1)$. Let $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ be the system corresponding to two qubits. Consider the normalized EPR state*

$$\psi = \frac{1}{\sqrt{2}}(e_1 \otimes e_1 + e_2 \otimes e_2) \in \mathcal{H}.$$

Let f be the self-adjoint operator on $\mathcal{H}_i$ given by the matrix $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Suppose that the qubit of $\mathcal{H}_1$ is held by Alice and the qubit of $\mathcal{H}_2$ is held by Bob, and they are located far apart. If Alice measure the observable f on her qubit, then what she actually does is measuring $f \otimes I$ of ψ. Now $f \otimes I$ has two eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 0$, with eigenspaces $\mathcal{H}_f(\lambda_1) = span\{e_1\} \otimes \mathcal{H}_2$ and $\mathcal{H}_f(\lambda_2) = span\{e_2\} \otimes \mathcal{H}_2$. The projections p_1 and p_2 are given by $p_1 = e_1^T \cdot e_1 \otimes I$ and $p_2 = e_2^T \cdot e_2 \otimes I$. It follows that for both $i = 1, 2$, the probability of collapsing ψ to $e_i \otimes e_i$ is $(\psi, p_i\psi) = \frac{1}{2}$. It follows that if Bob measures f on his qubit immediately after Alice, the result will be identical to Alice's. This is somewhat disturbing, because it means that Alice was able to convey the value of her qubit to Bob, essentially instantaneously.

Example 3.6 (GHZ Pseudo-telepathy Game)**.**

Example 3.7 (Bell's Inequality). *For an angle α let A_α be the observable on $\mathcal{H} = \mathbf{C}^2 \otimes \mathbf{C}^2$ given by*

$$A_\alpha = \begin{bmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix}.$$

A_α has eigenvalues ± 1. Let

$$\phi = \frac{1}{\sqrt{2}}(e_1 \otimes e_2 - e_2 \otimes e_1) \in \mathcal{H}$$

and let $X_{\alpha\beta}$ be the outcome of $A_\alpha \otimes A_\beta$ on ϕ. $X_{\alpha\beta}$ is a ± 1 valued random variable whose expected value is

$$E[X_{\alpha\beta}] = (\phi, (A_\alpha \otimes A_\beta)\phi) =$$

$$\frac{1}{2}(e_1 \otimes e_2 - e_2 \otimes e_1, A_\alpha e_1 \otimes A_\beta e_2 - A_\alpha e_2 \otimes A_\beta e_1) =$$

$$-\cos(\alpha - \beta).$$

Let $p_+(\alpha,\beta) = \Pr[X_{\alpha\beta} = 1]$ and $p_-(\alpha,\beta) = \Pr[X_{\alpha\beta} = -1]$. Then $p_+(\alpha,\beta) + p_-(\alpha,\beta) = 1$ and by the above $p_+(\alpha,\beta) - p_-(\alpha,\beta) = -\cos(\alpha-\beta)$, hence

$$p_-(\alpha,\beta) = \frac{1 + \cos(\alpha - \beta)}{2}.$$

Theorem 3.8 (Bell). *There do not exist random variables Y^1_θ, Y^2_θ with ± 1 values such that $X_{\alpha\beta} = Y^1_\alpha \cdot Y^2_\beta$ for all α, β.*

Proof. *Suppose to the contrary that $X_{\alpha\beta} = Y^1_\alpha \cdot Y^2_\beta$ for any α, β. In particular*

$$p_-(\alpha,\beta) = \Pr[Y^1_\alpha \neq Y^2_\beta]$$

and

$$\Pr[Y^1_\theta \neq Y^2_\theta] = \Pr[X_{\theta\theta} = -1] = 1.$$

It follows that for any α, β, γ

$$\Pr[Y^1_\alpha \neq Y^2_\beta] + \Pr[Y^1_\beta \neq Y^2_\gamma] + \Pr[Y^1_\gamma \neq Y^2_\alpha] =$$

$$\Pr[Y^1_\alpha = Y^1_\beta] + \Pr[Y^1_\beta = Y^1_\gamma] + \Pr[Y^1_\gamma = Y^1_\alpha] \geq 1.$$

but choosing $\alpha = 0, \beta = \frac{2\pi}{3}, \gamma = \frac{4\pi}{3}$ we obtain

$$\Pr[Y^1_\alpha \neq Y^2_\beta] + \Pr[Y^1_\beta \neq Y^2_\gamma] + \Pr[Y^1_\gamma \neq Y^2_\alpha] = 3 \cdot \frac{1 + \cos\frac{2\pi}{3}}{2} = \frac{3}{4},$$

a contradiction.

□

4 Spacetimes

In the following subsections we briefly describe three notions of spacetimes: Newton's, Galilei's and Minkowski's.

4.1 Newton Spacetime

Let V be a real linear space and let E be a set. Suppose that the additive group of V acts on E on the right, and denote the action of $v \in V$ on $p \in E$ by $p + v$.

Definition 4.1. *The pair* $\mathbb{A} = (E, V)$ *is an* affine space *if the action is simple transitive, i.e. for any* $p, p' \in E$ *there exists a unique* $v \overset{def}{=} p' - p \in V$ *such that* $p + v = p'$. *The dimension of* $\mathbb{A}$ *is* $\dim V$.

Definition 4.2. Newton $(1, d)$-Spacetime *is a 4-tuple* $(\mathbb{A}, \mathfrak{t}, \tau, h)$ *where* $\mathbb{A} = (E, V)$ *is an affine* $(d+1)$*-space,* $\mathfrak{t} \in V$, $\tau \in V^*$ *such that* $\tau(\mathfrak{t}) = 1$ *and* $h(\cdot, \cdot)$ *is an inner product on* $S = \tau^{-1}(0)$. *Let* $p : V \to S$ *be the projection corresponding to the direct sum decomposition* $V = span\{\mathfrak{t}\} \oplus S$. *Let* $O(S)$ *denote the orthogonal group of* S *with respect to* $h(\cdot, \cdot)$.

Definition 4.3. *An* automorphism *of* $(\mathbb{A}, \mathfrak{t}, \tau, h)$ *is a bijective map* $f : \mathbb{A} \to \mathbb{A}$ *such that the following conditions holds:*

(i) For any $\mathfrak{a} \in \mathbb{A}, v \in V$, *the vector* $\gamma_f(v) = f(\mathfrak{a} + v) - f(\mathfrak{a})$ *is independent of* $\mathfrak{a}$, *and the mapping* γ_f *is an element of* $GL(V)$.

(ii) $\gamma_f(\mathfrak{t}) = \mathfrak{t}$.

(iii) The restriction of γ_f *to* S *is an element of* $O(S)$.

Definition 4.4. *The* Newton Group $\mathcal{N}(V)$ *is the semidirect product* $O(S) \ltimes V$, *where the action of the orthogonal group* $O(S)$ *on* V *is given by*

$$\phi(v) = v - p(v) + \phi(p(v)).$$

Note that

$$\dim \mathcal{N}(V) = \dim O(S) + \dim V = \binom{\dim V - 1}{2} + d + 1 = \frac{d^2 + d + 2}{2}.$$

Fix an element $\mathfrak{o} \in \mathbb{A}$. An element $\eta = (\phi, u) \in O(S) \ltimes V = \mathcal{N}(V)$ gives rise to a map $f_\eta : \mathbb{A} \to \mathbb{A}$ given by

$$f_\eta(\mathfrak{a}) = \mathfrak{o} + (\phi(p(\mathfrak{a} - \mathfrak{o})) + \tau(\mathfrak{a} - \mathfrak{o})\mathfrak{t} + u)\,. \tag{58}$$

Claim 4.5. *(i)* $f_\eta \in \mathrm{Aut}\,((\mathbb{A}, \mathfrak{t}, \tau, h))$. *(ii) Any element of* $\mathrm{Aut}\,((\mathbb{A}, \mathfrak{t}, \tau, h))$ *is of the form* f_η *for some* $\theta \in \mathcal{N}(V)$.

Proof. (i) Let $\mathfrak{a} \in \mathbb{A}$ and $v \in V$. By (69)

$$\begin{aligned}\gamma_{f_\eta}(v) &= f_\eta(\mathfrak{a}+v) - f_\eta(\mathfrak{a}) \\ &= (\phi(p(\mathfrak{a}+v-\mathfrak{o})) + \tau(\mathfrak{a}+v-\mathfrak{o})\mathfrak{t}) - (\phi(p(\mathfrak{a}-\mathfrak{o})) + \tau(\mathfrak{a}-\mathfrak{o})\mathfrak{t}) \\ &= (\phi(p(\mathfrak{a}+v-\mathfrak{o})) - \phi(p(\mathfrak{a}-\mathfrak{o}))) + (\tau(\mathfrak{a}+v-\mathfrak{o})\mathfrak{t} - \tau(\mathfrak{a}-\mathfrak{o})\mathfrak{t}) \\ &= \phi(p(v)) + \tau(v)\mathfrak{t}\end{aligned} \tag{59}$$

It readily follows from (59) that $\gamma_{f_\eta}(\mathfrak{t}) = \mathfrak{t}$ and that the restriction of γ_{f_η} to S is $\phi \in O(S)$. (ii) Exercise.

□

Definition 4.6. *A* Newtonian Reference Frame *(abbreviated $\mathcal{N}$-frame) is a pair $(\mathfrak{o}, B)$, where $\mathfrak{o} \in \mathbb{A}$ and $B = [e_1, \ldots, e_d]$ is an ordered h-orthonormal basis of S. The* coordinates *assigned to an event $\mathfrak{a} \in \mathbb{A}$ by $(\mathfrak{o}, B)$ is the vector $(x_0, \ldots, x_d) \in \mathbf{R}^{d+1}$ where $\mathfrak{a} = \mathfrak{o} + x_0\mathfrak{t} + \sum_{i=1}^d x_i e_i$. x_0 is the* time *coordinate of $\mathfrak{a}$ and $(x_1, \ldots, x_d)$ are the* spatial *coordinates of $\mathfrak{o}$. An* Newtonian Inertial Observer *(abbreviated $\mathcal{N}$-observer) is a parametrized line $\gamma : \mathbf{R} \to \mathbb{A}$ given by $\gamma(\theta) = \mathfrak{o} + \theta(\mathfrak{t} + v)$ where $v \in S$.*

Let $(\mathfrak{o}, B), (\mathfrak{o}', B')$ be two $\mathcal{N}$-frames, where $B = [e_1, \ldots, e_d]$ and $B' = [e'_1, \ldots, e'_d]$. Let $T \in O(d)$ denote the transition matrix between B and B', i.e. $[e_1, \ldots, e_d]T = [e'_1, \ldots, e'_d]$. Let $z = (z_0, \ldots, z_d)$ such that $\mathfrak{o} - \mathfrak{o}' = z_0\mathfrak{t} + \sum_{i=1}^d z_i e_i$. Let $x = (x_0, \ldots, x_d)$ be the coordinates of $\mathfrak{a}$ in $(\mathfrak{o}, B)$ and let $x' = (x'_0, \ldots, x'_d)$ be the coordinates of $\mathfrak{a}$ in $(\mathfrak{o}', B')$.

Claim 4.7.

$$x' = \begin{bmatrix} 1 & 0 \\ 0 & T^{-1} \end{bmatrix} (x + z). \tag{60}$$

Proof.

$$\mathfrak{a} = \mathfrak{o} + x_0\mathfrak{t} + \sum_{i=1}^d x_i e_i = \mathfrak{o}' + x'_0\mathfrak{t} + \sum_{i=1}^d x'_i e'_i.$$

therefore

$$\begin{aligned}&(x_0 + z_0)\mathfrak{t} + \begin{bmatrix} e_1 & \cdots & e_d \end{bmatrix} \begin{bmatrix} x_1 + z_1 \\ \vdots \\ x_d + z_d \end{bmatrix} \\ &= x'_0\mathfrak{t} + \begin{bmatrix} e_1 & \cdots & e_d \end{bmatrix} T \begin{bmatrix} x'_1 \\ \vdots \\ x'_d \end{bmatrix}.\end{aligned} \tag{61}$$

Apply τ to (64) we obtain $x'_0 = x_0 + z_0$ and

$$T^{-1} \begin{bmatrix} x_1 + z_1 \\ \vdots \\ x_d + z_d \end{bmatrix}.$$

□

Definition 4.8. *A* Newtonian Particle *(abbreviated $\mathcal{N}$-particle) of mass $m > 0$ is a pair (m, γ) where $\gamma : (a, b) \to \mathbb{A}$ is differentiable and satisfies $\tau(\dot\gamma(\theta)) = 1$ for all θ. The following are some mechanical attributes of the particle:* Velocity *is $v = p(\dot\gamma)$,* Acceleration *is $a = p(\ddot\gamma)$,* Momentum*is $= mv$,* Force *on the particle is $F = \frac{d(m\dot\gamma(\theta))}{d\theta}$,* Kinetic Energy *of the particle is $KE = \frac{1}{2}m|v|^2$. The particle is* inertial *if $\ddot\gamma(\theta) = 0$, i.e. if there exist $\mathfrak{o} \in \mathbb{A}$ and $v \in S$ such that $\gamma(\theta) = \mathfrak{o} + \theta(\mathfrak{t} + v)$ for all θ.*

Remarks. *(i) In Newton spacetime, the time interval between two events $\mathfrak{a}, \mathfrak{b} \in \mathbb{A}$ is independent of the reference frame. Indeed, let $\mathfrak{a}, \mathfrak{b} \in \mathbb{A}$, and let $(\mathfrak{o}, B), (\mathfrak{o}', B')$ be two $\mathcal{N}$-frames. Let $x = (x_0, \ldots, x_d), y = (y_0, \ldots, y_d)$ be respectively the coordinates of $\mathfrak{a}, \mathfrak{b}$ according to $(\mathfrak{o}, B)$. Let $x' = (x'_0, \ldots, x'_d), y' = (y'_0, \ldots, y'_d)$ be respectively the coordinates of $\mathfrak{a}, \mathfrak{b}$ according to $(\mathfrak{o}', B')$. By (60)*

$$y' - x' = \begin{bmatrix} 1 & 0 \\ 0 & T^{-1} \end{bmatrix} (y - x)$$

and therefore $y'_0 - x'_0 = y_0 - x_0$.
(ii) An $\mathcal{N}$-particle (m, γ) is at rest *with respect to a $\mathcal{N}$-frame $(\mathfrak{o}, B)$, if the coordinates vector $x(\theta) = (x_0(\theta), \ldots, x_d(\theta))$ of $\gamma(\theta)$ in $(\mathfrak{o}, B)$ satisfies $(x_1(\theta), \ldots, x_d(\theta)) = (c_1, \ldots, c_d)$ for all θ. Let $y(\theta) = (y_0(\theta), \ldots, y_d(\theta))$ be the coordinates vector of $\gamma(\theta)$ relative to another $\mathcal{N}$-frame $(\mathfrak{o}', B')$ Then by (60)*

$$\begin{bmatrix} y_1(\theta) \\ \vdots \\ y_d(\theta) \end{bmatrix} = T^{-1} \begin{bmatrix} x_1(\theta) + z_1 \\ \vdots \\ x_d(\theta) + z_t \end{bmatrix} = T^{-1} \begin{bmatrix} c_1 + z_1 \\ \vdots \\ c_d + z_d \end{bmatrix},$$

and so the particle is at rest also relative to $(\mathfrak{o}', B')$. Thus, in Newton spacetime there is an absolute notion of rest.

4.2 Galilei Spacetime

Definition 4.9. Galilei $(1, d)$-Spacetime *is a 3-tuple $(\mathbb{A}, \tau, h)$ where $\mathbb{A} = (E, V)$ is an affine $(d + 1)$-space, $0 \neq \tau \in V^*$, and $h(\cdot, \cdot)$ is an inner product on $S = \tau^{-1}(0)$.*

Definition 4.10. *An* automorphism *of $(\mathbb{A}, \tau, h)$ is a bijective map $f : \mathbb{A} \to \mathbb{A}$ such that the following conditions hold:*

(i) For any $\mathfrak{a} \in \mathbb{A}, v \in V$, the vector $\gamma_f(v) = f(\mathfrak{a} + v) - f(\mathfrak{a})$ is independent of $\mathfrak{a}$, and is an element of $GL(V)$.

(ii) $\tau(\gamma_f(v)) = \tau(v)$ for all $v \in V$.

(iii) The restriction of γ_f to S is an element of $O(S)$.

Let

$$\tilde{\mathcal{G}}(V) = \left\{\phi \in GL(V) : \tau\phi = \tau, \phi_{|S} \in O(S)\right\}.$$

Fixing a basis $[e_0, \ldots, e_d]$ where $[e_1, \ldots, e_d]$ is an orthonormal basis of S and $\tau(e_0) = 1$, the matrices representing elements of $\tilde{\mathcal{G}}(V)$ are of the form $\begin{bmatrix} 1 & 0 \\ \alpha & T \end{bmatrix}$ where $\alpha \in \mathbf{R}^d$ and $T \in O(d)$.

Definition 4.11. *The* Galilei Group $\mathcal{G}(V)$ *is the semidirect product* $\tilde{\mathcal{G}}(V) \ltimes V$ *with the natural action of* $\tilde{\mathcal{G}}(V)$ *on* V*.*

Note that

$$\dim \mathcal{G}(V) = \dim \tilde{\mathcal{G}}(V) + \dim V = (d + \binom{d}{2}) + (d+1) = \binom{d+2}{2}.$$

Fix an element $\mathfrak{o} \in \mathbb{A}$. An element $\eta = (\phi, u) \in \tilde{\mathcal{G}}(V) \ltimes V = \mathcal{G}(V)$ gives rise to a map $f_\eta : \mathbb{A} \to \mathbb{A}$ given by

$$f_\eta(\mathfrak{a}) = \mathfrak{o} + \phi(\mathfrak{a} - \mathfrak{o}) + u. \tag{62}$$

Claim 4.12. *(i)* $f_\eta \in \mathrm{Aut}\,((\mathbb{A}, \tau, h))$*. (ii) Any element of* $\mathrm{Aut}\,((\mathbb{A}, \tau, h))$ *is of the form* f_η *for some* $\theta \in \mathcal{G}(V)$*.*

Proof. (i) Similar to the proof of Claim 4.5(i). (ii) Exercise.

□

Definition 4.13. *A* Galilien Reference Frame *(abbreviated $\mathcal{G}$-frame) is a pair* $(\mathfrak{o}, B)$*, where* $\mathfrak{o} \in \mathbb{A}$ *and* $B = [e_0, \ldots, e_d]$ *is an ordered basis of* V *such that* $\tau(e_0) = 1$ *and* $[e_1, \ldots, e_d]$ *is an orthnormal basis of* S*. The* coordinates *assigned to an event* $\mathfrak{a} \in \mathbb{A}$ *by* $(\mathfrak{o}, B)$ *is the vector* $(x_0, \ldots, x_d) \in \mathbf{R}^{d+1}$ *where* $\mathfrak{a} = \mathfrak{o} + \sum_{i=0}^d x_i e_i$*.* x_0 *is the* time *coordinate of* $\mathfrak{a}$ *and* $(x_1, \ldots, x_d)$ *are the* spatial *coordinates of* $\mathfrak{o}$*.*

Remark. *An $\mathcal{G}$-frame can be equivalently specified by giving an inertial particle together with an orthonormal basis of* S*.*

Let $(\mathfrak{o}, B), (\mathfrak{o}', B')$ be two $\mathcal{G}$-frames, where $B = [e_0, e_1, \ldots, e_d]$ and $B' = [e'_0, e'_1, \ldots, e'_d]$. Let $M \in GL(\mathbf{R}^{d+1})$ denote the transition matrix between B and B', i.e. $[e_0, \ldots, e_d]M = [e'_0, \ldots, e'_d]$. Note that M is of the form $M = \begin{bmatrix} 1 & 0 \\ \alpha & T \end{bmatrix}$ where $\alpha \in \mathbf{R}^d$ and $T \in O(S)$. Let $z = (z_0, \ldots, z_d)$ such that $\mathfrak{o} - \mathfrak{o}' = \sum_{j=0}^d z_j e_j$. Let $x = (x_0, \ldots, x_d)$ be the coordinates of $\mathfrak{a}$ in $(\mathfrak{o}, B)$ and let $x' = (x'_0, \ldots, x'_d)$ be the coordinates of $\mathfrak{a}$ in $(\mathfrak{o}', B')$.

Claim 4.14.

$$x' = M^{-1}(x + z) = \begin{bmatrix} 1 & 0 \\ -T^{-1}\alpha & T^{-1} \end{bmatrix}(x + z). \tag{63}$$

Proof.

$$\mathfrak{a} = \mathfrak{o} + \sum_{i=0}^{d} x_i e_i = \mathfrak{o}' + \sum_{i=0}^{d} x_i' e_i'.$$

therefore

$$\begin{bmatrix} e_0 & \cdots & e_d \end{bmatrix} \begin{bmatrix} x_0 + z_0 \\ \vdots \\ x_d + z_d \end{bmatrix} = \begin{bmatrix} e_0 & \cdots & e_d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \alpha & T \end{bmatrix} \begin{bmatrix} x_0' \\ \vdots \\ x_d' \end{bmatrix}, \tag{64}$$

implying (63).

□

Definition 4.15. *A* Galilien Particle *(abbreviated $\mathcal{G}$-particle) of mass $m > 0$ is a pair (m, γ) where $\gamma : (a, b) \to \mathbb{A}$ is differentiable and satisfies $\tau(\dot{\gamma}(\theta)) = 1$ for all t. The particle is* inertial *if $\ddot{\gamma}(\theta) = 0$, i.e. if there exist $\mathfrak{o} \in \mathbb{A}$ and $u \in \tau^{-1}(1)$ such that $\gamma(\theta) = \mathfrak{o} + \theta u$ for all θ.*

Remarks. *(i) In Galilei spacetime, the time interval between two events $\mathfrak{a}, \mathfrak{b} \in \mathbb{A}$ is independent of the reference frame. Indeed, let $\mathfrak{a}, \mathfrak{b} \in \mathbb{A}$, and let $(\mathfrak{o}, B), (\mathfrak{o}', B')$ be two $\mathcal{N}$-frames. Let $x = (x_0, \ldots, x_d), y = (y_0, \ldots, y_d)$ be respectively the coordinates of $\mathfrak{a}, \mathfrak{b}$ according to $(\mathfrak{o}, B)$. Let $x' = (x_0', \ldots, x_d'), y' = (y_0', \ldots, y_d')$ be respectively the coordinates of $\mathfrak{a}, \mathfrak{b}$ according to $(\mathfrak{o}', B')$. By (63)*

$$y' - x' = \begin{bmatrix} 1 & 0 \\ -T^{-1}\alpha & T^{-1} \end{bmatrix} (y - x)$$

and therefore $y_0' - x_0' = y_0 - x_0$. This also implies that if two events $\mathfrak{a}, \mathfrak{b}$ are simultaneous, i.e. $\tau(\mathfrak{b} - \mathfrak{a}) = 0$, then their spatial distance does not depend on the reference frame.
(ii) Let (m, γ) be a $\mathcal{G}$-particle with coordinates $(x_0(\theta), \ldots, x_d(\theta))$ in a $\mathcal{G}$-frame $(\mathfrak{o}, B)$. Then $x_0(\theta) = \lambda + \theta$ for some constant λ.
(iii) A $\mathcal{G}$-particle (m, γ) is at rest *with respect to a $\mathcal{G}$-frame $(\mathfrak{o}, B)$, if the coordinates vector $x(\theta) = (x_0(\theta), \ldots, x_d(\theta))$ of $\gamma(\theta)$ in $(\mathfrak{o}, B)$ satisfies $(x_1(\theta), \ldots, x_d(\theta)) = (c_1, \ldots, c_d)$ for all θ. In contrast with the Newtonian case, in Galilei spacetime there is no notion of absolute rest. Indeed, suppose (m, γ) is at rest relative to $(\mathfrak{o}, B)$ where $B = [e_0, \ldots, e_d]$ and let $B' = [e_0', \ldots, e_d']$ where $e_0' = e_0 + u$ for some $0 \neq u \in S$ and $e_i' = e_i$ for $1 \leq i \leq d$. Let $u = \sum_{i=1}^{d} \alpha_i e_i$. Then the transition matrix between B and B' is $M = \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix}$, where $\alpha = (\alpha_1, \ldots, \alpha_d)^T$. Let $y(\theta) = (y_0(\theta), \ldots, y_d(\theta))$ be the*

coordinates vector of $\gamma(\theta)$ relative to the $\mathcal{N}$-frame $(\mathfrak{o}, B')$. Then by (63)

$$y(\theta) = M^{-1}x(\theta) = \begin{bmatrix} 1 & 0 \\ -\alpha & 1 \end{bmatrix} x(\theta) = \begin{bmatrix} x_0(\theta) \\ x_1(\theta) + \alpha_1 x_0(\theta) \\ \vdots \\ x_d(\theta) + \alpha_d x_0(\theta) \end{bmatrix}$$

$$= \begin{bmatrix} \lambda + \theta \\ (c_1 + \alpha_1\lambda) + \alpha_1\theta \\ \vdots \\ (c_d + \alpha_d\lambda) + \alpha_d\theta \end{bmatrix} = \begin{bmatrix} \lambda \\ c_1 + \alpha_1\lambda \\ \vdots \\ c_d + \alpha_d\lambda \end{bmatrix} + \theta \begin{bmatrix} 1 \\ \alpha_1 \\ \vdots \\ \alpha_d \end{bmatrix}.$$

4.3 Minkowski Spacetime

We first discuss Lorentz vector spaces. Let V be an n-dimensional real vector space with a scalar product, i.e. symmetric bilinear nondegenerate form $\langle \cdot, \cdot \rangle$. For $v \in V$ let $|v| = |\langle v, v\rangle|^{1/2}$. A basis $e_1, \ldots, e_n$ of V is orthonormal if $\langle e_i, e_j \rangle = \epsilon_i \delta_{ij}$ where $\epsilon_i \in \{\pm 1\}$.

Claim 4.16. *(i) Any orthonormal set $\{e_1, \ldots, e_k\}$ can be extended to an orthonormal basis of V. (ii) The* index $|\{1 \leq i \leq n : \epsilon_i = -1\}|$ *is independent of the orthonormal basis. (iii) $v = \sum_{i=1}^{n} \epsilon_i \langle v, e_i \rangle e_i$ for any vector $v \in V$.*

Definition 4.17. *A* Lorentz vector space *is a real vector space of dimension $n \geq 2$ with scalar product $\langle \cdot, \cdot \rangle$ of index 1. The vectors $v \in V$ are classified as follows:*

- *Timelike: $\langle v, v \rangle < 0$.*
- *Null: $v \neq 0$ and $\langle v, v \rangle = 0$.*
- *Spacelike: $\langle v, v \rangle > 0$ or $v = 0$.*

Claim 4.18. *If $v \in V$ is timelike, then $v^\perp$ is spacelike, $V = \mathbf{R}v \oplus v^\perp$, and the restriction of $\langle \cdot, \cdot \rangle$ to $v^\perp$ is an inner product.*

Let $\mathcal{F} = \{u \in V : \langle u, u \rangle < 0\}$ be the set of timelike vectors. Define a relation $\sim$ on $\mathcal{F}$ by $v \sim w$ if $\langle v, w \rangle < 0$.

Claim 4.19. *$\sim$ is an equivalence relation.*

Proof. Reflexivity and symmetry are clear. For transitivity, assume that $u \sim v$ and $u \sim w$. We may assume that $|u| = 1$. write $v = au + v'$ and $w = bu + w'$ where $v', w' \in u^\perp$. $\langle u, v \rangle < 0$ implies that $a > 0$ and $\langle u, w \rangle < 0$ implies that $b > 0$. Moreover $\langle v, v \rangle < 0$ implies that $a^2 > \langle v', v' \rangle$ and $\langle w, w \rangle < 0$ implies that $b^2 > \langle w', w' \rangle$. As $u^\perp$ is an inner product space it follows that $|\langle v', w' \rangle| \leq |v'| \cdot |w'| \leq |ab| = ab$. Therefore

$$\langle v, w \rangle < -ab + \langle v', w' \rangle \leq |\langle v', w' \rangle| - |ab| < 0.$$

□

For $u \in \mathcal{F}$ let $C(u)$ denote the $\sim$-equivalence class of u. Clearly $C(-u) = -C(u)$ and $C(u) \cup C(-u) = \mathcal{F}$.

Claim 4.20. *(i) Reverse Cauchy-Schwarz inequality:*

$$|\langle u, v\rangle| \geq |u| \cdot |v| \tag{65}$$

for any $u, v \in \mathcal{F}$, with equality iff $u = \lambda v$. (ii) Reverse triangle inequality: If $u, v, w \in V$ satisfy $v - u, w - v \in \mathcal{F}$ and $C(v-u) = C(w-v)$, then $|w-u| \geq |v-u| + |w-v|$. (iii) If $C(u) = C(v)$ then there exists a unique number $\alpha \geq 0$, called the hyperbolic angle *between u and v) such that*

$$-\frac{\langle u, v\rangle}{|u| \cdot |v|} = \cosh \alpha. \tag{66}$$

(iv) $C(u)$ is an open convex cone.

Proof. (i) Write $v = au + v'$ where $v' \in u^{\perp}$. Then $0 > \langle v, v\rangle = a^2\langle u, u\rangle + |v'|^2$. Hence

$$\langle u, v\rangle^2 = a^2\langle u, u\rangle^2 = \langle u, u\rangle \left(\langle v, v\rangle - |v'|^2\right) \geq \langle u, u\rangle \cdot \langle v, v\rangle.$$

(ii) Let $u_1 = v - u$ and $v_1 = w - v$. Then

$$\begin{aligned}-|w-u|^2 &= \langle w-u, w-u\rangle = \langle u_1 + v_1, u_1 + v_1\rangle \\ &= \langle u_1, u_1\rangle + 2\langle u_1, v_1\rangle + \langle v_1, v_1\rangle = -\left(|u_1|^2 + 2|\langle u_1, v_1\rangle| + |v_1|^2\right) \\ &\leq -\left(|u_1|^2 + 2|u_1| \cdot |v_1| + |v_1|^2\right) \\ &= -\left(|u_1| + |v_1|\right)^2.\end{aligned}$$

(iii) The map $\alpha \to \cosh \alpha$ maps $\mathbf{R}_{\geq 0}$ injectively onto $\mathbf{R}_{\geq 1}$. Hence, if $C(u) = C(v)$ then (66) follows from the reverse Cauchy-Schwarz inequality $-\frac{\langle u,v\rangle}{|u|\cdot|v|} \geq 1$.
(iv) Let $v, w \in C(u)$ and $\alpha, \beta > 0$. Then

$$\begin{aligned}\langle \alpha u + \beta v, \alpha u + \beta v\rangle &= \alpha^2\langle u, u\rangle + 2\alpha\beta\langle u, v\rangle + \beta^2\langle v, v\rangle \\ &\leq \alpha^2\langle u, u\rangle - 2\alpha\beta|u| \cdot |v| + \beta^2\langle v, v\rangle \\ &= -\left(\alpha^2|u|^2 + 2\alpha\beta|u| \cdot |v| + \beta^2|v|\right)^2 \\ &= -(\alpha|u| + \beta|v|)^2 < 0.\end{aligned}$$

□

Definition 4.21. *Let V be a Lorentz vector space. The* Lorentz Group *of V is*

$$L(V) = \{g \in GL(V) : \langle gu, gv\rangle = \langle u, v\rangle \text{ for all } u, v \in V\}$$

The Restricted Lorentz Group *of V is the connected component $L^0(V)$ of the identity element of L, i.e.*

$$L^0(V) = \{g \in L(V) : \det g = 1\ ,\ gC(v) = C(v) \text{ for } v \in \mathcal{F}\}.$$

Let D be the diagonal matrix $diag(-1,1,\ldots,1) \in GL(\mathbf{R}^{d+1})$.

$$O(1,d) = \{A \in GL(\mathbf{R}^{d+1}) : A^t D A = D\}$$

and

$$O(1,d)^0 = \{A \in O(1,d) : \det A = 1, A_{00} \geq 1\}. Let$$

Let $B = [e_0, \ldots, e_d]$ be an orthonormal basis of V. Let $g \in GL(V)$ an let A be the matrix representing g according to the basis B.

Claim 4.22. *(i) $g \in L(V)$ iff $A \in O(1,d)$. (ii) $g \in L^0(V)$ iff $O(1,d)^0$.*

Proof. Let $a_i = (a_{0i}, \ldots, a_{di})^t$ be the i-th column of A. Then

$$\begin{aligned}\langle ge_k, ge_\ell \rangle &= \left\langle \sum_{i=0}^d a_{ik} e_i, \sum_{j=0}^d a_{j\ell} e_j \right\rangle = \sum_{i,j} a_{ik} a_{j\ell} \langle e_i, e_j \rangle \\ &= \sum_{i=0}^d a_{ik} a_{i\ell} = \left(A^t D A\right)_{k\ell}.\end{aligned} \tag{67}$$

Now $g \in L(V)$ iff $\langle ge_k, ge_\ell \rangle = \epsilon_k \delta_{k\ell}$ for all k, ℓ. On the other hand, $A \in O(1,d)$ iff $(A^t D A)_{k\ell} = D_{k\ell} = \epsilon_k \delta_{k\ell}$ for all k, ℓ. Hence (i) follows from (67). For (ii) note that $\langle ge_0, e_0 \rangle = -A_{00}$ and

$$-1 = \langle e_0, e_0 \rangle = \langle ge_0, ge_0 \rangle = -A_{00}^2 + \sum_{i=1}^d A_{i0}^2. \tag{68}$$

Therefore, if $g \in L^0(V)$ then $-A_{00} = \langle ge_0, e_0 \rangle < 0$, hence $A_{00} \geq 1$ by (68). The other direction is similar.

□

For $\alpha \in \mathbf{R}$ let

$$L(\alpha) = \begin{pmatrix} \cosh\alpha & \sinh\alpha \\ \sinh\alpha & \cosh\alpha \end{pmatrix} \in O(1,1)^0$$

and

$$\tilde{L}(\alpha) = \begin{pmatrix} L(\alpha) & 0 \\ 0 & I_{d-1} \end{pmatrix} \in O(1,d)^0.$$

Definition 4.23. Minkowski $(1,d)$-Spacetime *is a pair $(\mathbb{A}, \eta)$ where $\mathbb{A} = (E, V)$ is an affine $(d+1)$-space, and $\eta(u,v) = \langle u, v \rangle$ is a Lorentz scalar product on V. A* time orientation *on V is a choice of $v \in \mathcal{F}$ and the resulting* future cone *$C(v)$.*

In the following we fix $v \in \mathcal{F}$ and define the *future cone* by $C^+ = C(v)$.

Definition 4.24. *An* automorphism *of $(\mathbb{A}, \eta)$ is a bijective map $f : \mathbb{A} \to \mathbb{A}$ such that For any $\mathfrak{a} \in \mathbb{A}$, the map $\gamma_f : V \to V$ given by $\gamma_f(v) = f(\mathfrak{a} + v) - f(\mathfrak{a})$ is an element of $L^0(V)$.*

Definition 4.25. *The* Poincaré Group $\mathcal{P}(V)$ *is the semidirect product* $L^0(V) \ltimes V$ *with the natural action of* $L^0(V)$ *on* V.

Note that

$$\dim \mathcal{P}(V) = \dim L^0(V) + \dim V = \binom{d+1}{2} + (d+1) = \binom{d+2}{2}.$$

Fix an element $\mathfrak{o} \in \mathbb{A}$. An element $\lambda = (\phi, u) \in L^0(V) \ltimes V = \mathcal{P}(V)$ gives rise to a map $f_\lambda : \mathbb{A} \to \mathbb{A}$ given by

$$f_\lambda(\mathfrak{a}) = \mathfrak{o} + \phi(\mathfrak{a} - \mathfrak{o}) + u. \tag{69}$$

Claim 4.26. *(i)* $f_\lambda \in \mathrm{Aut}\,((\mathbb{A}, \eta))$. *(ii) Any element of* $\mathrm{Aut}\,(\mathbb{A}, \eta))$ *is of the form* f_λ *for some* $\lambda \in \mathcal{P}(V)$.

Proof. (i) Similar to the proof of Claim 4.5(i). (ii) Exercise.

□

Definition 4.27. *A* Special-Relativistic Reference Frame *(abbreviated* $\mathcal{R}$*-frame) is a pair* $(\mathfrak{o}, B)$*, where* $\mathfrak{o} \in \mathbb{A}$ *and* $B = [e_0, \ldots, e_d]$ *is an orthonormal basis of* V*, i.e.* $\langle e_i, e_j \rangle = \delta_{ij}\epsilon_i$ *where* $(\epsilon_0, \ldots, \epsilon_d) = (-1, 1, \ldots, 1)$*. We further assume that* $e_0 \in C^+$*. The* coordinates *assigned to an event* $\mathfrak{a} \in \mathbb{A}$ *by* $(\mathfrak{o}, B)$ *is the vector* $(x_0, \ldots, x_d) \in \mathbf{R}^{d+1}$ *where* $\mathfrak{a} = \mathfrak{o} + \sum_{i=0}^d x_i e_i$. x_0 *is the* time *coordinate of* $\mathfrak{a}$ *and* $(x_1, \ldots, x_d)$ *are the* spatial *coordinates of* $\mathfrak{a}$.

Definition 4.28. *A* Special-Relativistic Material Particle *(abbreviated* $\mathcal{R}$*-particle) of rest mass* $m_0 > 0$ *is a pair* (m_0, γ) *where* $\gamma : (a, b) \to \mathbb{A}$ *satisfies* $\dot\gamma(\theta) \in C^+$ *and* $\langle \dot\gamma(\theta), \dot\gamma(\theta) \rangle = -1$ *for all* θ*. The parameter* θ *is the* proper time *of the particle. The particle is* free *if* $\ddot\gamma(\theta) = 0$*, i.e. if there exist* $\mathfrak{o} \in \mathbb{A}$ *and* $\mathfrak{t} \in C^+$ *such that* $\langle \mathfrak{t}, \mathfrak{t} \rangle = -1$ *and* $\gamma(\theta) = \mathfrak{o} + \theta\mathfrak{t}$ *for all* θ*. Such* γ *is also called a* free observer*. A* Lightlike Particle *is a* $\gamma : (a, b) \to \mathbb{A}$ *such that* $\gamma(\theta) = \mathfrak{a} + \theta v$*, where* v *is lightlike and future directed, i.e.* $0 \neq v \in \overline{C^+} \setminus C^+$.

Let $x = (x_0, \ldots, x_d), y = (y_0, \ldots, y_d)$ be coordinates assigned to distinct events $\mathfrak{a}$, $\neq \mathfrak{b} \in \mathbb{A}$ respectively by a reference frame $(\mathfrak{o}, B)$.

Claim 4.29. *Let* $v = \mathfrak{b} - \mathfrak{a}$*. We consider the following cases.*
(i) v *is null. Then* $(y_0 - x_0)^2 = \sum_{i=1}^d (y_i - x_i)^2$*. Furthermore, then sign of* $y_0 - x_0$ *is independent of the frame. In physical terms: all frames agree that the events* $\mathfrak{a}, \mathfrak{b}$ *occur on a worldline of a lightlike particle, and on which of the events occurred first.*
(ii) v *is timelike. Then there exists a frame* $(\mathfrak{o}, B)$ *such that* $x_i = y_i$ *for all* $1 \leq i \leq d$*. In physical terms: There exists a frame for which* $\mathfrak{a}$ *and* $\mathfrak{b}$ *occupy the same spatial location.*
(iii) v *is spacelike. Then for any* $t_0 \in \mathbf{R}$ *there exists a frame* $(\mathfrak{o}, B)$ *such that* $y_0 - x_0 = t_0$*. In physical terms: Any number can be realized as the time separation of* $\mathfrak{a}$ *and* $\mathfrak{b}$.

Proof. Let $(\mathfrak{o}, B)$ be a reference frame where $B = [e_0, \ldots, e_d]$. Let $\mathfrak{a} = \mathfrak{o} + \sum_{i=0}^{d} x_i e_i$ and $\mathfrak{b} = \mathfrak{o} + \sum_{i=0}^{d} y_i e_i$. Then $v = \sum_{i=0}^{d}(y_i - x_i)e_i$.
(i) If v is null then

$$0 = \langle v, v \rangle = \left\langle \sum_{i=0}^{d}(y_i - x_i)e_i, \sum_{i=0}^{d}(y_i - x_i)e_i \right\rangle = (y_0 - x_0)^2 - \sum_{i=1}^{d}(y_i - x_i)^2.$$

Moreover, if v is future directed, i.e. $v \in \overline{C^+}$, then a limit argument shows that so is gv for every $g \in L^0(V)$. Hence the sign of $y_0 - x_0$ is independent of the reference frame.
(ii) Suppose v is timelike. Let $e_0 = \frac{v}{|v|}$, and complete e_0 to an orthonormal basis $B = [e_0, \ldots, e_d]$. Clearly $y - x = (|v|, 0, \ldots, 0)$.
(iii) Suppose v is spacelike. Let $e_1 = \frac{v}{|v|}$ and complete e_1 to an orthonormal basis $B = [e_0, \ldots, e_d]$. Then $y - x = (0, |v|, 0, \ldots, 0)$. Therefore $\tilde{L}(\alpha)(y - x) = |v|(\sinh\alpha, \cosh\alpha, 0, \ldots, 0)$, hence by choosing the appropriate reference frame, the time coordinate of $y - x$ can attain any real value.

□

Claim 4.30. *Let (m_0, γ) be a material particle $\gamma : (0, \tilde{\theta}) \to \mathbb{A}$, where $\gamma(0) = \mathfrak{a}$ and $\gamma(\tilde{\theta}) = b$. Then $v = \mathfrak{b} - \mathfrak{a} \in C^+$ and*

$$\tilde{\theta} \leq |v|. \tag{70}$$

Equality occurs iff $\gamma(\theta) = \mathfrak{a} + \theta\frac{v}{|v|}$.

Proof. As C^+ is a convex cone and $\dot{\gamma}(\theta) \in C^+$ for all $\theta \in [0, \tilde{\theta}]$, it follows that

$$v = \int_{\theta=0}^{\tilde{\theta}} \dot{\gamma}(\theta) d\theta \in C^+.$$

For any $\theta \in [0, \tilde{\theta}]$, the reverse Cauchy-Schwarz inequality (65) implies that

$$-\langle \dot{\gamma}(\theta), v \rangle = |\langle \dot{\gamma}(\theta), v \rangle| \geq |\dot{\gamma}(\theta)| \cdot |v| = |v|.$$

Hence

$$|v|^2 = -\langle v, v \rangle = -\int_{\theta=0}^{\tilde{\theta}} \langle \dot{\gamma}(\theta), v \rangle d\theta \geq \int_{\theta=0}^{\tilde{\theta}} |v| d\theta = |v| \cdot \tilde{\theta},$$

and therefore $\tilde{\theta} \leq |v|$.

□

Remark. *In physical terms, Claim 4.30 says that if* $\mathfrak{b} - \mathfrak{a} \in C^+$*, then among all material particles that experience both* $\mathfrak{a}$ *and* $\mathfrak{b}$*, the free particle will record the maximal proper time between the events. For example, fix a frame* $(\mathfrak{o}, B)$ *where* $B = [e_0, \dots, e_d]$*. Let* $0 \leq \lambda < 1$ *and consider the material particles* $(m_0, \gamma_1), (m_0, \gamma_2)$ *where*

$$\gamma_1(\theta) = \begin{cases} \mathfrak{o} + \theta \cdot \frac{e_0 + \lambda e_1}{\sqrt{1-\lambda^2}} & 0 \leq \theta \leq 1 \\ \mathfrak{o} + \frac{2\lambda e_1}{\sqrt{1-\lambda^2}} + \theta \cdot \frac{e_0 - \lambda e_1}{\sqrt{1-\lambda^2}} & 1 \leq \theta \leq 2, \end{cases}$$

and

$$\gamma_2(\theta) = \mathfrak{o} + \theta e_0$$

for $0 \leq \theta \leq \frac{2}{\sqrt{1-\lambda^2}}$*. Both particles experience* $\mathfrak{o}$ *and* $\mathfrak{o} + \frac{2e_0}{\sqrt{1-\lambda^2}}$*. The first particle's journey takes* 2 *time units, while the second particle's journey takes* $\frac{2e_0}{\sqrt{1-\lambda^2}}$*.*

Definition 4.31. *Let* (m_0, γ) *be an* $\mathcal{R}$*-particle. The relativistic* Velocity, Momentum, *and* Force *are given respectively by* $v(\theta) = \dot{\gamma}(\theta)$*,* $p(\theta) = m_0 \dot{\gamma}(\theta)$*, and* $f(\theta) = \frac{d(m_0 \dot{\gamma}(\theta))}{d\theta}$*.*

Let (m_0, γ) be an $\mathcal{R}$-particle. Given a reference frame $(\mathfrak{o}, B)$ where $B = [e_0, \dots, e_d]$ write $\gamma(\theta) = \mathfrak{o} + \sum_{i=0}^{d} x_i(\theta) e_i$. The classical velocity, momentum and force of the particle with respect to the frame are defined as follows. The classical velocity is

$$v_c(\theta) = \sum_{i=1}^{d} \frac{\dot{x}_i(\theta)}{\dot{x}_0(\theta)} e_i.$$

Note that $\dot{x}_0(\theta) = (1 - |v_c(\theta)|^2)^{-\frac{1}{2}}$, and that

$$v(\theta) = \dot{x}_0(\theta) \left(e_0 + v_c(\theta) \right).$$

Let $m(\theta) = m_0 \dot{x}_0(\theta)$. The classical momentum is

$$p_c(\theta) = m(\theta) v_c(\theta).$$

Thus

$$p(\theta) = m(\theta) e_0 + p_c(\theta).$$

The classical force is

$$f_c(\theta) = \frac{dp_c}{dx_0}.$$

Thus

$$f(\theta) = \dot{m}(\theta) e_0 + \dot{x}_0(\theta) f_c(\theta).$$

Now $\langle v(\theta), v(\theta) \rangle = -1$ implies that

$$\begin{aligned} 0 = \langle f(\theta), v(\theta) \rangle &= \langle \dot{m}(\theta) e_0 + \dot{x}_0(\theta) f_c(\theta), \dot{x}_0(\theta) \left(e_0 + v_c(\theta) \right) \rangle \\ &= \dot{x}_0(\theta) \left(-\dot{m}(\theta) + \dot{x}_0(\theta) \langle f_c(\theta), v_c(\theta) \rangle \right). \end{aligned}$$

It follows that $\dot{m}(\theta) = \dot{x}_0(\theta) \langle f_c(\theta), v_c(\theta) \rangle$. Therefore

$$f(\theta) = \dot{x}_0(\theta) \big(\langle f_c(\theta), v_c(\theta) \rangle e_0 + f_c(\theta) \big). \tag{71}$$

5 Differential Forms - A Brief Introduction

Let $\Delta_k = \mathrm{conv}\{\mathrm{e_0}, \dots, \mathrm{e_k}\}$ be the standard k-simplex, where $e_0 = 0$ and $e_1, \dots, e_k$ are the first k unit vectors in $\mathbf{R}^\infty$. For $0 \le i \le k+1$, let $\epsilon_i : \Delta_k \to \Delta_{k+1}$ be the affine map given by

$$\epsilon(e_j) = \begin{cases} e_j & 0 \le j < i, \\ e_{j+1} & i \le j \le k. \end{cases}$$

Let $I^k = [0,1]^k$, be the standard unit k-cube. For $\epsilon \in \{0,1\}, 1 \le i \le k+1$ define $\phi_{i,\epsilon} : I_k \to I_{k+1}$ be the affine map

$$\phi_{i,\epsilon}(x_1, \dots, x_k) = (x_1, \dots, x_{i-1}, \epsilon, x_i, \dots, x_k).$$

Let $M \subset \mathbf{R}^n$ be an open set. A singular k-simplex (k-cube) in M is a smooth map $T : \Delta_k \to M$ ($T : I_k \to M$). Let $S_k(M)$ ($C_k(M)$) denote respectively the free R-modules generated by the singular k-simplices (k-cubes). The differential $\partial_{k+1} : S_{k+1}(M) \to S_k(M)$ is given by

$$\partial_{k+1} T = \sum_{i=0}^{k+1} (-1)^i T \circ \epsilon_i.$$

The differential $\partial_{k+1} : C_{k+1}(M) \to C_k(M)$ is given by

$$\partial_{k+1} T = \sum_{i=0}^{k+1} \sum_{\epsilon=0}^{1} (-1)^{i+\epsilon} T \circ \phi_{i,\epsilon}.$$

Let U be an open set in $\mathbf{R}^k$. For a differential form

$$\omega = \sum_{I=\{i_1 < \dots < i_k\}} a_I(x_1, \dots, x_n) dx_{i_1} \wedge \dots \wedge dx_{i_k} \in \Omega^k(M)$$

and a smooth map $T = (T_1, \dots, T_n) : U \to M$, define

$$\int_T \omega = \sum_{I=\{i_1 < \dots < i_k\}} \int_{u \in U} a_I(T(u)) \frac{\partial(T_{i_1}, \dots, T_{i_k})}{\partial(u_1, \dots, u_k)}(u)\, du_1 \cdots du_k.$$

Theorem 5.1 (Change of Variables)**.** *Let $M \subset \mathbf{R}^n$, $N \subset \mathbf{R}^m$ be open and let $f : M \to N$ be a smooth map. Then for any $c \in C_k(M)$ and $\omega \in \Omega^k(N)$*

$$\int_{f_* c} \omega = \int_c f^* \omega. \tag{72}$$

Theorem 5.2 (Stokes Formula for singular cubical chains)**.** *For any $\omega \in \Omega^k(M)$ and $c \in C_{k+1}(M)$*

$$\int_c d_k \omega = \int_{\partial_{k+1} c} \omega. \tag{73}$$

Proof. It suffices to establish (77) for $M = \mathbf{R}^{k+1}$ and for $T = \mathrm{Id} \in \mathrm{C_{k+1}(M)}$, where $\mathrm{Id} : \mathrm{I^{k+1}} \to \mathrm{M}$ is the identity map $\mathrm{Id(x) = x}$. Indeed, if N is open in $\mathbf{R}^n$, $\omega \in \Omega^k(N)$ and $T : I^{k+1} \to N$, then

$$\begin{aligned}\int_T d_k\omega &= \int_{T\circ \mathrm{Id}} d_k\omega \\ &= \int_{\mathrm{Id}} T^*(d_k\omega) = \int_{\mathrm{Id}} d_k(T^*\omega) \\ &= \int_{\partial_{k+1}\mathrm{Id}} T^*\omega = \int_{T_*(\partial_{k+1}\mathrm{Id})} \omega \\ &= \int_{\partial_{k+1}(T_*\mathrm{Id})} \omega = \int_{\partial_{k+1}T} \omega.\end{aligned} \tag{74}$$

We now check the case $T = \mathrm{Id} : \mathrm{I^{k+1}} \to \mathrm{M} = \mathbf{R}^{\mathrm{k+1}}$ and $\omega \in \Omega^k(\mathbf{R}^k)$. By linearity, we may assume that $\omega = a(x_1, \ldots, x_{k+1})dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_{k+1}$. On one hand

$$\begin{aligned}&\int_{\mathrm{Id}} d_k\omega = \int_{\mathrm{Id}} \sum_{j=1}^{k+1} \frac{\partial a}{\partial x_j} dx_j \wedge dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_{k+1} \\ &= (-1)^{i-1} \int_{\mathrm{Id}} \frac{\partial a}{\partial x_i} dx_1 \wedge \cdots \widehat{dx_{k+1}} \\ &= (-1)^i \sum_{\epsilon=0}^{1} \int_{u=(u_1,\ldots,u_k)\in I^k} (-1)^\epsilon a(u_1, \ldots, u_{i-1}, \epsilon, u_i, \ldots, u_k) du_1 \cdots du_k.\end{aligned} \tag{75}$$

On the other hand

$$\begin{aligned}&\int_{\partial_{k+1}\mathrm{Id}} \omega = \sum_{j=1}^{k+1} \sum_{\epsilon=0}^{1} (-1)^{j+\epsilon} \int_{\phi_{j,\epsilon}} \omega \\ &= \sum_{j=1}^{k+1} \sum_{\epsilon=0}^{1} (-1)^{j+\epsilon} \int_{u\in I^k} a(\phi_{j,\epsilon}(u)) \frac{\partial \left((\phi_{j,\epsilon})_1, \ldots, \widehat{(\phi_{j,\epsilon})_i}, \ldots, (\phi_{j,\epsilon})_{k+1}\right)}{\partial (u_1, \ldots, u_k)} du_1 \cdots du_k \\ &= \sum_{\epsilon=0}^{1} (-1)^{i+\epsilon} \int_{u=(u_1,\ldots,u_k)\in I^k} a(u_1, \ldots, u_{i-1}, \epsilon, u_i, \ldots, u_k) du_1 \cdots du_k.\end{aligned} \tag{76}$$

Comparing (75) and (76) we obtain

$$\int_{\mathrm{Id}} d_k\omega = \int_{\partial_{k+1}\mathrm{Id}} \omega.$$

□

Corollary 5.3 (Stokes Formula for singular simplicial chains)**.** *For any $\omega \in \Omega^k(M)$ and $c \in S_{k+1}(M)$*

$$\int_c d_k\omega = \int_{\partial_{k+1}c} \omega. \tag{77}$$

Suppose $\langle \cdot, \cdot \rangle$ is a scalar product on an n-dimensional V, and let $B = [e_1, \ldots, e_n]$ be a fixed orthonormal basis of V. Let $dx_1, \ldots dx_n$ be the dual basis of $\Omega^n(V)$, and let $\omega = dx_1 \wedge \cdots \wedge dx_n$. Let M be open in V. Define the Hodge duality map $* = *_V : \Omega^k(M) \to \Omega^{n-k}(M)$ as the $C^\infty(M)$-linear map given on basis elements by $(dx_I) \wedge (dx_J) = \langle *dx_I, dx_J \rangle \tau$. For example, if V is the 4-dimensional Lorentz space, then $*(dx_0 \wedge dx_2) = -dx_1 \wedge dx_3$. The codifferential $d^* : \Omega^k(M) \to \Omega^{k-1}(M)$ is defined by $d^* = *d*$.
For basic k-form $\omega = dx_{i_1} \wedge \cdots \wedge dx_{i_k} \in \Omega^k(M)$ and a basic vector field $\lambda = \frac{\partial}{\partial x_j} \in TM$ let

$$\begin{aligned} i(\lambda)\omega &= \sum_{\ell=1}^{k} (-1)^{\ell+1} \left\langle \frac{\partial}{\partial x_j}, dx_{i_\ell} \right\rangle dx_{i_1} \wedge \cdots \wedge \widehat{dx_{i_\ell}} \wedge \cdots \wedge dx_{i_k} \\ &= \sum_{\ell=1}^{k} (-1)^{\ell+1} \delta_{j,i_\ell} \epsilon_j dx_{i_1} \wedge \cdots \wedge \widehat{dx_{i_\ell}} \wedge \cdots \wedge dx_{i_k}. \end{aligned}$$

Extend this definition by linearity over $C^\infty(M)$ to general forms $\omega \in \Omega^k(M)$, $\lambda \in TM$.

5.1 Maxwell Equations

Throughout this section we work in 4-Minkowsky space, i.e. $d = 4$. Let $B = [e_0, \ldots, e_3]$ be a fixed orthonormal basis of V and identify $x = (x_0, \ldots, x_3)$ with $\sum_{i=0}^{3} x_i e_i$. Consider a unit charge that travels with a path $\gamma(\theta)$. Let $f_c(\theta)$ denote the classical force acting on this test charge. The Lorentz force formula states that

$$f_c(\theta) = E(\gamma(\theta)) + v_c(\theta) \times B(\gamma(\theta)) \tag{78}$$

where $E(x) = \sum_{i=1}^{3} E_i(x) e_i$ is the electric field and and $B(x) = \sum_{i=1}^{3} B_i(x) e_i$ is the magnetic field. We next reconstruct the actual force $f(\theta)$. By (78)

$$\begin{aligned} \langle f_c(\theta), v_c(\theta) \rangle &= \langle E(\gamma(\theta)) + v_c(\theta) \times B(\gamma(\theta)), v_c(\theta) \rangle \\ &= \langle E(\gamma(\theta)), v_c(\theta) \rangle. \end{aligned} \tag{79}$$

Hence, by (71)

$$\begin{aligned} f(\theta) &= \dot{x}_0(\theta) \big(\langle f_c(\theta), v_c(\theta) \rangle e_0 + f_c(\theta) \big) \\ &= \dot{x}_0(\theta) \big(\langle E(\gamma(\theta)), v_c(\theta) \rangle e_0 + E(\gamma(\theta)) + v_c(\theta) \times B(\gamma(\theta)) \big) \end{aligned} \tag{80}$$

For $x \in V$ define a linear mapping $\tilde{\mathcal{F}} : V \to V$ as follows. For $v = (v_0, \ldots, v_3) \in V$ let

$$\tilde{\mathcal{F}}(x)(v) = \left(\sum_{i=1}^{3} E_i(x) v_i \right) e_0 + v_0 E(x) + \left(\sum_{i=1}^{3} v_i e_i \right) \times B(x). \tag{81}$$

Then

$$f(\theta) = \tilde{\mathcal{F}}(\gamma(\theta)) v(\theta). \tag{82}$$

Let $\mathcal{F}(\cdot,\cdot)$ be the bilinear form on $V\times V$ given by $\mathcal{F}(u,v)=\langle u,\tilde{\mathcal{F}}(x)v\rangle$. Let

$$F=F(x)=\begin{pmatrix} 0 & -E_1 & -E_2 & -E_3\\ E_1 & 0 & B_3 & -B_2\\ E_2 & -B_3 & 0 & B_1\\ E_3 & B_2 & -B_1 & 0\end{pmatrix}. \tag{83}$$

Then

$$\begin{aligned}&\mathcal{F}(u,v)=\langle u,\tilde{\mathcal{F}}(x)v\rangle\\ &=\left\langle u,\left(\sum_{i=1}^{3}v_iE_i, v_0E_1+v_2B_3-v_3B_2, v_0E_2+v_3B_1-v_1B_3, v_0E_3+v_1B_2-v_2B_1\right)\right\rangle\\ &=u^tFv.\end{aligned}$$

Let $S=\operatorname{span}\{e_1,e_2,e_3\}$. Note that for $1\le j\le 3$, the Hodge duality operator $*_S$ satisfies $*_Sdx_j=dx_k\wedge dx_\ell$, where (j,k,ℓ) are a cyclic shift of $(1,2,3)$. In the sequel, we identify $\mathcal{F}$ with the 2-form

$$\frac{1}{2}\sum_{k,\ell}F_{k,\ell}dx_k\wedge dx_\ell=\sum_{j=1}^{3}E_jdx_j\wedge dx_0+\sum_{j=1}^{3}B_j(*_Sdx_j).$$

$\mathcal{F}$ is called the *Electromagnetic Tensor*. The assumption that the Lorentz formula holds in each reference frame implies that $\mathcal{F}$ is globally defined, namely if $\mathcal{F}'$ is the 2-form constructed according to the reference frame $[e_0',\dots,e_3']$, then $\mathcal{F}=\mathcal{F}'$.

Claim 5.4.

$$d\mathcal{F}=\sum_{j=1}^{3}\left(\frac{\partial B}{\partial x_0}+\nabla\times E\right)_j(*_Sdx_j)\wedge dx_0+(\text{div }B)dx_1\wedge dx_2\wedge dx_3. \tag{84}$$

$$d^*\mathcal{F}=\sum_{j=1}^{3}\left((\nabla\times B)_j-\frac{\partial E_j}{\partial x_0}\right)dx_j-(\text{div }E)dx_0. \tag{85}$$

Proof. Note that $*dx_1\wedge dx_0=-dx_2\wedge dx_3$, $*dx_2\wedge dx_0=-dx_3\wedge dx_1$ and $*dx_3\wedge dx_0=-dx_1\wedge dx_2$. Furthermore $*dx_2\wedge dx_3=dx_1\wedge dx_0$, $*dx_3\wedge dx_1=dx_2\wedge dx_0$ and $*dx_1\wedge dx_2=dx_3\wedge dx_0$. hence

$$*\mathcal{F}=-(E_1x_2\wedge dx_3+E_2dx_3\wedge dx_1+E_3dx_1\wedge dx_2)+\sum_{i=1}^{3}B_idx_i\wedge dx_0.$$

□

Let J denote the electric current vector field and let ρ denote the charge density. The Maxwell's equations are the following relations between E, B, J and ρ.

$$\operatorname{div} B = 0, \tag{86}$$

$$\nabla \times E = -\frac{\partial B}{\partial t}, \tag{87}$$

$$\operatorname{div} E = 4\pi\rho, \tag{88}$$

$$\nabla \times B = \frac{\partial E}{\partial t} + 4\pi J. \tag{89}$$

Remarks. *Eq. (86) asserts that there are no magnetic charges. Eq. (87) is Faraday's Law. Eq. (88) is Gauss Law. Eq. (89) is Ampere-Maxwell Law.*

Let

$$\mathcal{J}^b = -\rho\, dx_0 + \sum_{k=1}^{3} J_k\, dx_k.$$

Claim 5.4 implies that

$$\left(\operatorname{div} B = 0 \ \ \& \ \ \nabla \times E = -\frac{\partial B}{\partial t}\right) \quad \Longleftrightarrow \quad d\mathcal{F} = 0$$

and

$$\left(\operatorname{div} E = 4\pi\rho \ \ \& \ \ \nabla \times B = \frac{\partial E}{\partial t} + 4\pi J\right) \quad \Longleftrightarrow \quad d^*\mathcal{F} = 4\pi\mathcal{J}^b.$$

Using the fact that the 2-form $\mathcal{F}$ is globally defined, we will now show that (86) implies (87), and that (88) implies (89). We start with the first case. A flat M in a Lorentz space V is spacelike if it is a translate of a spacelike linear subspace of V.

Proposition 5.5. *Let M be an open set in an n-dimensional Lorentz vector space V. Let $0 \le k \le n-1$ and let $0 \ne \omega \in \Omega^k(M)$. Then there exists a spacelike hyperplane $H \subset V$ such that the inclusion $i_{H\cap M} : H \cap M \to M$ satisfies $i^*_{H\cap M}\omega \ne 0$.*

Proof. Choose an orthonormal basis $[e_1, \ldots, e_n]$ of V and let $\omega(x) = \sum_{I \in \binom{[n]}{k}} a_I(x) dx_I$. We consider two cases:
(i) $a_I(x) \ne 0$ for some $I \in \binom{[n]\setminus\{1\}}{k}$. Let $H = p + \operatorname{span}\{e_i\}_{i=2}^n$ and define ϕ from a small neighborhood of $0 \in \mathbf{R}^{n-1}$ to M by $\phi(y_2, \ldots, y_n) = p + \sum_{i=2}^n y_i e_i$. Then H is spacelike and

$$\phi^* i^*_{M\cap H}\omega = \phi^*\omega = \sum_{I \in \binom{[n]\setminus\{1\}}{k}} a_I \left(p + \sum_{i=2}^n y_i e_i\right) dy_I \ne 0.$$

(ii) $a_I(x) = 0$ for all $I \in \binom{[n]\setminus\{1\}}{k}$. Then there exist $I_0 = \{1 = i_1, i_2 \ldots, i_k\}$ and $p \in M$ such that $a_{I_0}(p) \neq 0$. Let $m \notin I_0$ and let $H = p + \mathrm{span}\{e_1 + 2e_m, e_2, \ldots, \widehat{e_m}, \ldots, e_n\}$. Define ϕ from a small neighborhood of $0 \in \mathbf{R}^{n-1}$ to M by

$$\phi(y_1, \ldots, \widehat{y_m}, \ldots, y_n) = p + y_1(e_1 + 2e_m) + \sum_{2 \leq i \neq m} y_i e_i.$$

Then

$$\phi^* i^*_{M\cap H}\omega = \phi^*\omega = \sum_{1\in I} a_I \left(p + y_1(e_1 + 2e_m) + \sum_{2\leq i\neq m} y_i e_i \right) dy_I \neq 0.$$

□

Corollary 5.6. *If* $\mathrm{div}\, B = 0$ *in every reference frame, then* $d\mathcal{F} = 0$ *and hence* $\nabla \times E = -\frac{\partial B}{\partial x_0}$ *in every reference frame.*

Proof. Let M be a spacelike hyperplane in V and write $M = w+U$ where $w \in V$ and U is a spacelike linear hyperplane. Let $[e_0, e_1, e_2, e_3]$ be a reference frame of V such that $U = \mathrm{span}\{e_1, e_2, e_3\}$. Then there exists a $T \in \mathbf{R}$ such that $M = Te_0 + \mathrm{span}\{e_1, e_2, e_3\}$. Let (x_0, x_1, x_2, x_3) be the coordinates corresponding to the above frame. Eq. (84) implies that $i^*_M d\mathcal{F} = \mathrm{div}\, B\, dx_1 \wedge dx_2 \wedge dx_3 = 0$. Proposition 5.7 now implies that $d\mathcal{F} = 0$.

□

We now show that (88) implies (89).

Proposition 5.7. *Let* M *be an open subset of an* n*-dimensional Lorentz vector space* V*. Let* $1 \leq k \leq n$ *and let* $0 \neq \omega \in \Omega^k(V)$*. Then there exists a timelike vector* v *such that* $i(v)\omega \neq 0$*.*

Proof. Choose an orthonormal basis $[e_1, \ldots, e_n]$ of V and let $\omega(x) = \sum_{I\in\binom{[n]}{k}} a_I(x) dx_I$. We consider two cases:
(i) $a_I(x) \neq 0$ for some $1 \in I \in \binom{[n]}{k}$. Then

$$i(e_1)\omega = -\sum_{1\in I} a_I dx_{I\setminus\{1\}} \neq 0.$$

(ii) $a_I(x) = 0$ for all $I \in \binom{[n]}{k}$ such that $1 \in I$. Choose $I_0 = \{i_1, i_2 \ldots, i_k\} \subset [n] \setminus \{1\}$ and $p \in M$ such that $a_{I_0}(p) \neq 0$. Then $2e_1 + 2e_{i_1}$ is timelike and

$$i(2e_1 + e_{i_1})\omega = i(e_{i_1})\omega = \sum_{I\ni i_1} a_I(x) i(e_{i_1}) dx_I \neq 0.$$

□

Corollary 5.8. *If* $\operatorname{div} E = 4\pi\rho$ *in every reference frame, then* $d^*\mathcal{F} = 4\pi\mathcal{J}^b$ *and hence* $\nabla \times B = \frac{\partial E}{\partial x_0} + 4\pi J$ *in every reference frame.*

Proof. Let $0 \neq v$ be a timelike vector. Let $[e_0, e_1, e_2, e_3]$ be a frame such that $e_0 = \frac{v}{|v|}$. Eq. (85) implies that

$$d^*\mathcal{F} - 4\pi\mathcal{J}^b = \sum_{j=1}^{3} \left((\nabla \times B)_j - \frac{\partial E_j}{\partial x_0} - 4\pi J_j \right) dx_j - (\operatorname{div} E - 4\pi\rho) dx_0$$

The assumption that $\operatorname{div} E = 4\pi\rho$ in this frame implies that

$$i(v)\left(d^*\mathcal{F} - 4\pi\mathcal{J}^b\right) = -|v|(\operatorname{div} E - 4\pi\rho) = 0. \tag{90}$$

As (90) holds for any timelike v, it follows by Proposition 5.7 that $d^*\mathcal{F} = 4\pi\mathcal{J}^b$ and therefore also

$$\nabla \times B = \frac{\partial E}{\partial x_0} + 4\pi J.$$

□

6 Relativistic Wave Equations

As we saw earlier, a closed quantum system is associated with a Hilbert space $\mathcal{H}$, where the points of the projective space $P(\mathcal{H})$ are in one to one correspondence with the states of the system. An automorphism of the system is a bijective mapping $\varphi : P(\mathcal{H}) \to P(\mathcal{H})$ such that if $0 \neq u, v \in \mathcal{H}$ and $u' \in \varphi([u])$, $v' \in \varphi([v])$, then

$$\frac{|(u,v)|}{|u| \cdot |v|} = \frac{|(u',v')|}{|u'| \cdot |v'|}.$$

Theorem 6.1 (Wigner)**.** *Any such automorphism* φ *is of the form* $\varphi([u]) = [Au]$ *where* $A : \mathcal{H} \to \mathcal{H}$ *is either unitary or anti-unitary.*

Let $\tau : G \to \operatorname{Aut}(P(\mathcal{H}))$ be a representation of a connected Lie group G on $P(\mathcal{H})$. Theorem 6.1 implies that there exists a mapping $\rho : G \to U(\mathcal{H})$ and a function $\alpha : G \times G \to \mathbf{R}$ such that $\tau(g)([u]) = [\rho(g)(u)]$ and

$$\rho(g_1 g_2) = e^{i\alpha(g_1,g_2)} \rho(g_1)\rho(g_2). \tag{91}$$

for all $g_1, g_2 \in G$. Such ρ is called a *projective unitary representation* of G on $\mathcal{H}$. A quantum system $\mathcal{H}$ is an *elementary relativistic free particle*, if $\mathcal{H}$ is an irreducible projective unitary representation of G, i.e. if there are no $0 \neq \mathcal{H}_0 \subsetneq \mathcal{H}$ that are invariant under G. At this point we should specify G and study its projective representations. As a first step, let $G = \mathcal{P}$ be the Poincaré group, i.e. the group of symmetries of spacetime. We first recall the definition of $\mathcal{P}$ and some of its properties.

6.1 Geometry of the Lorentz and Poincaré Groups

In the present chapter we switch the sign convention to $(+---)$. Thus *Minkowski space* is $M = R^{1,3}$ is $\mathbf{R}^4$ with the indefinite quadratic form $(+---)$ metric, i.e. if $x = (x_0, x_1, x_2, x_3), y = (y_0, y_1, y_2, y_3)$ then

$$\langle x, y\rangle = x_0 y_0 - \sum_{i=1}^{3} x_i y_i.$$

Let $\|x\|^2 = \langle x, x\rangle = x_0^2 - \sum_{i=1}^3 x_i^2$. Let $\eta = \mathrm{diag}(1, -1, -1, -1) \in GL(\mathbf{R}^4)$. The *Lorentz group* $L = O(1,3)$ is defined by

$$\begin{aligned} L &= \{B \in GL(\mathbf{R}^4) : \|Bx\|^2 = \|x\|^2 \text{ for all } x \in \mathbf{R}^4\} \\ &= \{B \in GL(\mathbf{R}^4) : B^T \eta B = \eta\}. \end{aligned} \tag{92}$$

The *Proper Lorentz Group* is the connected component of $I \in L$:

$$L^0 = \{B = (B_{ij})_{i,j=0}^3 \in L : B_{00} \geq 1, \det B = 1\}.$$

Claim 6.2.

(i) L^0 is diffeomorphic to $\mathbf{R}^3 \times SO(3)$.

(ii) $SO(3)$ is diffeomorphic to $\mathbf{RP}^3$.

(iii) $SL(2, \mathbf{C})$ is diffeomorphic to $S^3 \times \mathbf{R}^3$.

Proof. (i) Let $H = \{x = (x_0, \ldots, x_3) \in \mathbf{R}^{1,3} : \langle x, x\rangle = 1\ ,\ x_0 \geq 1\}$. The map $x \to (x_1, x_2, x_3)$ is a diffeomorphism of H and $\mathbf{R}^3$. Define $\pi : L^0 \to H$ by $\pi(B) = Be_0$. Then π is onto: any $v_0 \in H$ can be completed to an orthnormal basis v_0, v_1, v_2, v_3 such that $B = [v_0, v_1, v_2, v_3] \in L^0$. Then

$$\pi^{-1}(v_0) = \left\{ B \cdot \begin{bmatrix} 1 & 0 \\ 0 & A \end{bmatrix} : A \in SO(3) \right\}.$$

Hence L^0 is an $SO(3)$-bundle over $H \cong \mathbf{R}^3$. As H is contractible it follows that L^0 is diffeomorphic to $SO(3) \times \mathbf{R}^3$.
(ii) Let $B^3(\pi) \subset \mathbf{R}^3$ be the closed ball of radius π. Define $f : B^3(\pi) \to SO(3)$, by $f(0) = I$ and for $0 \neq u \in B^3(\pi)$ let $f(u)$ be the rotation with angle $|u|$ around the ray $\mathbf{R}^+ \cdot u$. It is clear that f is injective on the interior of $B^3(\pi)$, while $f(u) = f(-u)$ for $|u| = \pi$. Thus f induces a diffeomorphism $\mathbf{RP}^3 \to SO(3)$.
(iii) The map $\pi : SL(2, \mathbf{C}) \to \mathbf{C}^2 - \{(0, 0)\}$ given by

$$\pi\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a \\ c \end{bmatrix}$$

is a PFB with fibre $\mathbf{C}$. It has a global section $s : \mathbf{C}^2 - \{(0,0\} \to SL(2,\mathbf{C})$ given by

$$s\left(\begin{bmatrix} a \\ c \end{bmatrix}\right) = \begin{bmatrix} a & -\frac{\bar{c}}{\gamma} \\ c & \frac{\bar{a}}{\gamma} \end{bmatrix}$$

where $\gamma = |a|^2 + |c|^2$. It follows that

$$SL(2,\mathbf{C}) \cong \big(\mathbf{C}^2 - \{(0,0)\}\big) \times \mathbf{C} \cong S^3 \times \mathbf{R}^3.$$

□

Let $H(2,\mathbf{C})$ denote the space of complex Hermitian 2×2 matrices. We identify $\mathbf{R}^{1,3}$ with $H(2,\mathbf{C})$ via the map

$$x = (x_0, x_1, x_2, x_3) \to \underline{x} = \begin{bmatrix} x_0 + x_3 & x_1 - ix_2 \\ x_1 + ix_2 & x_0 - x_3 \end{bmatrix}.$$

Note that $\det \underline{x} = \langle x, x \rangle$. Define a homomorphism $\phi : SL(2,\mathbf{C}) \to GL(H(2,\mathbf{C}))$ by

$$\phi(A)(\underline{x}) = A\underline{x}A^*.$$

Then for any $A \in SL(2,\mathbf{C})$ and $x \in H(2,\mathbf{C})$

$$\langle \phi(A)(\underline{x}), \phi(A)(\underline{x}) \rangle = \det \phi(A)(\underline{x}) = \det (A\underline{x}A^*) = \det \underline{x} = \langle x, x \rangle.$$

This, together with the connectivity of $SL(2,\mathbf{C})$, imply that $\phi(A) \in L^0$.

Proposition 6.3.

(a) ϕ *maps* $SL(2,\mathbf{C})$ *onto* L^0, *with* $\ker \phi = \{\pm I\}$.

(b) ϕ *maps* $SU(2) \subset SL(2,\mathbf{C})$ *onto* $SO(3) \subset L^0$ *with* $\ker \phi = \{\pm I\}$.

Remark: By Claim 6.2, $SL(2,\mathbf{C}) \cong S^3 \times \mathbf{R}^3$. Hence $SL(2,\mathbf{C})$ is the universal cover of L^0. Similarly, $SU(2) \cong S^3$ is the universal cover of $SO(3) \cong \mathbf{RP}^3$.

The *Poincaré Group* is the semidirect product $\mathcal{P} = L^0 \ltimes \mathbf{R}^{1,3}$. As defined earlier, an elementary relativistic free particle is an irreducible projective representation of of $\mathcal{P}$. In the next sections we will study some of these representations.

6.2 Tempered Distributions - a Brief Summary

Definition 6.4. *The* Schwartz Space $\mathcal{S}(\mathbf{R}^n)$ *consists of all* C^∞ *complex valued functions* f *on* $\mathbf{R}^n$ *such that* $\|x^\beta D^\alpha f\|_\infty < \infty$ *for all* α, β. *For example* $P(x)e^{-a|x|^2} \in \mathcal{S}(\mathbf{R}^n)$ *for any polynomial* $P(x)$ *and* $a > 0$. *A* Tempered Distribution *is an element of* $\mathcal{S}'(\mathbf{R}^n)$, *the space of continuous linear functionals on* $\mathcal{S}(\mathbf{R}^n)$. *I.e. a linear map* $T : \mathcal{S}(\mathbf{R}^n) \to \mathbf{C}$ *such that there exist* m, n *and* C *that satisfy*

$$|T(f)| \leq C\, Sup_{|\alpha| \leq m, |\beta| \leq n} \|x^\beta D^\alpha f\|_\infty$$

for all $f \in \mathcal{S}(\mathbf{R}^n)$.

Examples.
1. If g *is a polynomially bounded measurable function, i.e.* $(1 + |x|^2)^{-N} g(x) \in L^1(\mathbf{R}^n)$ *for some* N, *then the functional* T_g *given by* $T_g(f) = \int_x f(x)g(x)dx$ *is in* $\mathcal{S}'(\mathbf{R}^n)$.
2. If μ *is a measure on* $\mathbf{R}^n$ *such that* $\int_x (1 + |x|^2)^{-N} d\mu(x) < \infty$ *for some* N, *then the functional* T_μ *given by* $T_\mu(f) = \int_x f(x)d\mu(x)$ *is in* $\mathcal{S}'(\mathbf{R}^n)$.

Definition 6.5. *Let* $T \in \mathcal{S}'(\mathbf{R}^n)$ *and* $1 \leq k \leq n$. *The* differential $\partial_k T$ *is given by* $\partial_k T(f) = -T(\partial_k f)$.

Examples.
1. If $g \in \mathcal{S}(\mathbf{R}^n)$ *then* $\partial_k T_g = T_{\partial_k g}$.
2. Let $f = 1_{[0,\infty)}$ *be the Heaviside function. Then* $\partial T_f = \delta_0$.

Let $\langle \cdot, \cdot \rangle$ be the scalar product in $\mathbf{R}^n$ given by $\langle x, y \rangle = \sum_{k=1}^n \epsilon_k x_k y_k$.

Definition 6.6. *The* Fourier Transform *of* $f \in \mathcal{S}(\mathbf{R}^n)$ *is the function* $\mathcal{F}(f) \in \mathcal{S}(\mathbf{R}^n)$ *given by*

$$\mathcal{F}(f)(p) = \frac{1}{(2\pi)^{n/2}} \int_x f(x) e^{-i\langle p, x \rangle} dx.$$

Claim 6.7. *Let* $f \in \mathcal{S}(\mathbf{R}^n)$. *Then*

$$\begin{aligned} \mathcal{F}\left(\frac{\partial f}{\partial x^\alpha}\right)(p) &= (-i)^{|\alpha|} \epsilon^\alpha p^\alpha \mathcal{F}(f)(p), \\ \frac{\partial \mathcal{F}(f)}{\partial p^\beta}(p) &= (-i)^{|\beta|} \epsilon^\beta \mathcal{F}(x^\beta f)(p). \end{aligned} \tag{93}$$

Definition 6.8. *The* Fourier Transform *of* $T \in \mathcal{S}'(\mathbf{R}^n)$ *is the map on* $\mathcal{S}(\mathbf{R}^n)$ *given by* $\mathcal{F}(T)(g) = T(\mathcal{F}(g))$ *for all* $g \in \mathcal{S}(\mathbf{R}^n)$.

Claim 6.9. *If* $T \in \mathcal{S}'(\mathbf{R}^n)$ *then* $\mathcal{F}(T) \in \mathcal{S}'(\mathbf{R}^n)$.

6.3 Projective Representations of $\mathcal{P}$

Let $G = SL(2,\mathbf{C}) \ltimes \mathbf{R}^{1,3}$. It follows from Proposition 6.3 that G is a universal cover of $\mathcal{P}$.

Theorem 6.10 (Bargmann-Wigner)**.** *Let $\eta : \mathcal{P} \to U(\mathcal{H})$ be a unitary projective representation. Then there exists a unitary representation $\rho : G \to U(\mathcal{H})$ such that $[\rho(g)u] = \eta(\phi(g))[u]$ for all $g \in G$ and $0 \neq u \in \mathcal{H}$.*

The notion of induced representation described in Subsection 1.6 for finite groups, can be extended to certain families of infinite groups. Let G be a locally compact second countable group, and let H be a closed subgroup. Let $\lambda : H \to U(W)$ be unitary representation of H. Suppose that $X = G/H$ carries a G-invariant measure μ. Let $\overline{g} = gH \in X$. Let $C_{H,\lambda}(G,W)$ denote the space of Borel maps $s : G \to W$ such that

$$s(gh) = \lambda(h^{-1})s(g) \tag{94}$$

for all $g \in G, h \in H$ and such that

$$\|s\|^2 = \int_{x=gH\in X} |s(g)|^2 d\mu(x) < \infty. \tag{95}$$

Note that (94) and the unitarity of λ imply that $|s(g_1)| = |s(g_2)|$ if $g_1H = g_2H$, hence the integral in (95) is well defined. The *Induced Representation* $\eta = \mathrm{Ind}_H^G\lambda : G \to U\left(C_{H,\lambda}(G,W)\right)$ is given by $\eta(g)s(g') = s(g^{-1}g')$ for all $g, g' \in G$. If W is finite dimensional then $C_{H,\lambda}(G,W)$ can be identified with a space of sections of a certain vector bundle as follows. Let $\sim$ be the equivalence relation on $G \times W$ given by $(g,w) \sim (gh, \lambda(h^{-1})w)$ for all $(g,h,w) \in G \times H \times W$. Let W_λ be the quotient space $(G \times W)/\sim$. Denote by $[x,w]$ the equivalence class of $(x,w) \in G \times W$. The projection map $[g,w] \to [g] = gH$ defines a vector bundle over $X = G/H$. Define an action of G on G/H and on W_λ by $g(xH) = gxH$. Define an action of G on $G \times W$ by $g(x,w) = (gx,w)$. Clearly, if $(x_1,w_1) \sim (x_2,w_2)$, then $g(x_1,w_1) \sim g(x_2,w_2)$. Thus we get an action of G on W_λ. This action gives rise to a representation ρ of G on the space $\Gamma(W_\lambda)$ of square integrable sections of W_λ, given on a $\sigma \in \Gamma(W_\lambda)$ by $\rho(g)\sigma(x) = g\sigma(g^{-1}x)$. $T : C_{H,\lambda}(G,W) \to \Gamma(W_\lambda)$ by $T\phi([x]) = [x, \phi(x)]$. Note that T is well defined, i.e. if $[x_1] = [x_2]$ then $x_2 = x_1h$ for some $h \in H$ and thus

$$(x_2, \phi(x_2)) = (x_1h, \phi(x_1h)) = (x_1h, \lambda(h^{-1})\phi(x_1))$$

and hence $(x_2,\phi(x_2)) \sim (x_1,\phi(x_1))$.

Claim 6.11. *T is an isomorphism and the following diagram commutes:*

$$\begin{array}{ccc} C_{H,\lambda}(G,W) & \xrightarrow{T} & \Gamma(W_\lambda) \\ {\scriptstyle \eta(g)}\downarrow & & \downarrow{\scriptstyle \rho(g)} \\ C_{H,\lambda}(G,W) & \xrightarrow{T} & \Gamma(W_\lambda) \end{array}$$

Proof. Let $\phi \in C_{H,\lambda}(G,W)$ and let $x \in G$. Then

$$\begin{aligned}\rho(g)T\phi([x]) &= g(T\phi(g^{-1}[x])) \\ &= g[g^{-1}x, \phi(g^{-1}x)] = [x, \phi(g^{-1}x)] \\ &= [x, \eta(g)\phi(x)] = T\left(\eta(g)\phi\right)[x]\end{aligned}$$

□

In view of Claim 6.11 we will identify $\Gamma(W_\lambda)$ with $C_{H,\lambda}(G,W)$ and the representation η with $\rho = \mathrm{Ind}_H^G \lambda$.
We now describe a situation where the space of sections $\Gamma(W_\lambda)$ can be replaced by a simpler space. Suppose G acts on a space X and $H = \mathrm{Stab}_G(x_0)$ for some $x_0 \in X$. Let $\lambda : H \to U(W)$ is a unitary representation of H and suppose there exists a (not necessarily unitary) representation $\tau : G \to GL(W)$ such that $\tau(h) = \lambda(h)$ for all $h \in H$. Define $\psi : W_\lambda \to X \times W$ by

$$\psi\left([g,w]\right) = (gx_0, \tau(g)w). \tag{96}$$

Note that ψ is well defined vector bundle isomorphism. Recall that G acts on W_λ by $g'[g,w] = [g'g,w]$. Define an action of G on $X \times W$ by $g'(x,w) = (g'x, \tau(g')w)$. Then ψ is G-equivariant. we will identify $\Gamma(M \times W)$ with the space $C(M,W)$ of all tempered W-valued distributions on M. Let $\Psi : \Gamma(W_\lambda) = C_{H,\lambda}(G,W) \to \Gamma(M \times W)$ be given by $\Psi\phi(ax_0) = \tau(a)\phi(a)$. Note that Ψ is well defined: if $a'x_0 = ax_0$ then $a' = ah$ for some $h \in H$ and therefore

$$\tau(a')\phi(a') = \tau(ah)\phi(ah) = (\tau(a)\lambda(h))\left(\lambda(h^{-1})\phi(a)\right) = \tau(a)\phi(a).$$

The above reduction will play a key role in deriving the Dirac equation in Subsection 6.4.

The method of constructing the irreducible representations of semi-direct products, described in subsection 1.7 for the case of finite groups extends to certain Lie groups, in particular to $G = H \ltimes N$, where $H = SL(2,\mathbf{C})$ and $N = \mathbf{R}^{1,3}$. The character group $\widehat{\mathbf{R}^{1,3}}$ consists of all continuous maps $\mathbf{R}^{1,3} \to C^*$ and will be identified with $\mathbf{R}^{1,3}$ as follows: An element $p \in \mathbf{R}^{1,3}$ gives rise to the character χ_p given by $\chi_p(x) = \exp(i\langle p, x\rangle)$. It follows that the action of $SL(2,\mathbf{C})$ is given by $A(\chi_p) = \chi_{A(\underline{p})}$. We next compute the orbits of $SL(2,\mathbf{C})$ on $\mathbf{R}^{1,3}$, and the corresponding stabilizers. Let $m^2 = \langle p, p\rangle = \det \underline{p}$. Clearly, any orbit is contained in a level set of m^2. For example, fix $m > 0$ and consider the orbit of

$$\underline{p} = \begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix}.$$

Then A is in the stabilizer of $\underline{p}$ iff

$$A \begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix} A^* = \begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix},$$

i.e. iff $AA^* = I$, namely $A \in SU(2)$. In general we have the following orbits and their stabilizers.

Orbit	Representative	Stabilizer
$X_m^+ = \{p \in \mathbf{R}^{1,3} : \langle p, p\rangle = m^2 > 0\ ,\ p_0 > 0\}$	$\begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix}$	$SU(2)$
$X_m^+ = \{p \in \mathbf{R}^{1,3} : \langle p, p\rangle = m^2 > 0\ ,\ p_0 < 0\}$	$\begin{bmatrix} -m & 0 \\ 0 & -m \end{bmatrix}$	$SU(2)$
$X_0^+ = \{p \in \mathbf{R}^{1,3} : \langle p, p\rangle = 0\ ,\ p_0 > 0\}$	$\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$	$\tilde{E}(2)$
$X_0^- = \{p \in \mathbf{R}^{1,3} : \langle p, p\rangle = 0\ ,\ p_0 < 0\}$	$\begin{bmatrix} -2 & 0 \\ 0 & 0 \end{bmatrix}$	$\tilde{E}(2)$
$\{0\}$	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$	$SL_2(\mathbf{C})$
$Y_m = \{p \in \mathbf{R}^{1,3} : \langle p, p\rangle = -m^2 > 0\ \}$	$\begin{bmatrix} 0 & i\lvert m\rvert \\ i\lvert m\rvert & 0 \end{bmatrix}$	$SL_2(\mathbf{R})$

where

$$\tilde{E}(2) = \left\{ \begin{bmatrix} e^{i\theta} & b \\ 0 & e^{-i\theta} \end{bmatrix} \ :\ \theta \in \mathbf{R}, b \in \mathbf{C} \right\}.$$

6.4 Massive Particles

In this section we consider representations that arise from the orbits X_m^+, where $m > 0$. The stabilizer $K = SU(2)$ has for each $s \in \{0, \frac{1}{2}, 1, \frac{3}{2}, \ldots\}$ a $2s+1$-dimensional representation on the space V_s of homogenous polynomials in $\mathbf{C}[z_1, z_2]$ of degree $2s$. It follows that for each half integer s and $m > 0$ we get an irreducible unitary representation of G, where m corresponds to the rest mass of the particle, and s to its spin. We first construct an invariant measure μ_m^+ on X_m^+. Fix $a > 0$. For $\alpha > a$ let

$$D_\alpha = \{p = (p_0, \mathbf{p}) \in \mathbf{R} \times \mathbf{R}^3 : a < p_0^2 - \|\mathbf{p}\|^2 < \alpha, p_0 > 0\}.$$

For $f \in \mathcal{S}(\mathbf{R}^{1,3})$ let

$$F(f, \alpha) = \int_{p \in D_\alpha} f(p) dp$$

and let $J_\alpha(f) = \frac{dF(f,\alpha)}{d\alpha}$. Clearly $F(Af, \alpha) = F(f, \alpha)$ and hence

$$J_\alpha(Af) = J_\alpha(f) \tag{97}$$

for all $A \in L^0$. Let $\varphi : [a, \alpha] \times \mathbf{R}^3 \to D_\alpha$ be defined by

$$\varphi(u, \mathbf{p}) = \left((u + \|\mathbf{p}\|^2)^{\frac{1}{2}}, \mathbf{p} \right).$$

Then

$$\begin{aligned} F(f, \alpha) &= \int_{u=a}^{\alpha} \int_{\mathbf{p} \in \mathbf{R}^3} f(\varphi(u, \mathbf{p})) J_\varphi(u, \mathbf{p}) d\mathbf{p}\, du \\ &= \int_{u=a}^{\alpha} \int_{\mathbf{p} \in \mathbf{R}^3} \frac{f\left((u + \|\mathbf{p}\|^2)^{\frac{1}{2}}, \mathbf{p} \right) d\mathbf{p}}{2(u + \|\mathbf{p}\|^2)^{\frac{1}{2}}}\, du. \end{aligned}$$

It follows that

$$J_\alpha(f) = \int_{\mathbf{p} \in \mathbf{R}^3} \frac{f\left((\alpha + \|\mathbf{p}\|^2)^{\frac{1}{2}}, \mathbf{p} \right) d\mathbf{p}}{2(\alpha + \|\mathbf{p}\|^2)^{\frac{1}{2}}}.$$

Fix $m > 0$. Viewing J_{m^2} as a positive linear functional on $\mathcal{S}(\mathbf{R}^{1,3})$, there exists a measure $d\mu_m^+$ on $\mathbf{R}^{1,3}$ such that

$$J_{m^2}(f) = \int_{p \in \mathbf{R}^{1,3}} f(p) d\mu_m^+(p).$$

Thus

$$\int_{p \in \mathbf{R}^{1,3}} f(p) d\mu_m^+(p) = \int_{\mathbf{p} \in \mathbf{R}^3} \frac{f\left((m^2 + \|\mathbf{p}\|^2)^{\frac{1}{2}}, \mathbf{p} \right) d\mathbf{p}}{2(m^2 + \|\mathbf{p}\|^2)^{\frac{1}{2}}}. \tag{98}$$

Eq. (98) and (97) imply respectively that $d\mu_m^+$ is supported on X_m^+ and is L^0 invariant. We claim that the functional $T_{\mu_m^+}$ on $\mathcal{S}(\mathbf{R}^{1,3})$ is in $\mathcal{S}'(\mathbf{R}^{1,3})$. Indeed, (98) implies that $\int_{p \in \mathbf{R}^{1,3}} (1 + \sum_{i=0}^{3} p_i^2)^{-N} d\mu_m^+(p) < \infty$ for any $N > 5/2$.

Spin $s = 0$: The Klein-Gordon Equation

Here λ is the trivial representation and therefore the bundle W_λ is the trivial line bundle over X_m^+ and $\Gamma(W_\lambda) = L^2(X_m^+, \mu_m^+)$. We next obtain a more concrete realization of this representation. Let

$$\Box^2 = \frac{\partial^2}{\partial x_0^2} - \frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2} - \frac{\partial^2}{\partial x_3^2}.$$

Let $f \in L^2(X_m^+, \mu_m^+)$. Then $(\langle p, p \rangle - m^2) f d\mu_m^+ = 0$. Applying the Fourier transform and using (93), it follows that the distribution $\psi = \widehat{f d\mu_m^+}$ satisfies the Klein-Gordon equation

$$\left(\Box^2 + m^2 \right) \psi = 0. \tag{99}$$

Spin $s = \frac{1}{2}$: The Dirac Equation

Recall that the adjoint of a matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $\mathrm{adj}(A) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. Let

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad , \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad , \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Then $\sigma_k^2 = 1$ for all $1 \le k \le 3$ and $\sigma_k \sigma_\ell = -\sigma_\ell \sigma_k$ for all $1 \le k \ne \ell \le 3$. Define the Dirac matrices by

$$\gamma_0 = \begin{bmatrix} 0 & I_2 \\ I_2 & 0 \end{bmatrix} \quad \text{and} \quad \gamma_k = \begin{bmatrix} 0 & -\sigma_k \\ \sigma_k & 0 \end{bmatrix} \quad \text{for } 1 \le k \le 3.$$

Then $\gamma_k^2 = \epsilon_k$ for all $0 \le k \le 3$ and $\gamma_k \gamma_\ell = -\gamma_\ell \gamma_k$ for all $0 \le k \ne \ell \le 3$. Let $p = (p_0, p_1, p_2, p_3) \in \mathbf{R}^{1,3}$. For a matrix $P \in M_2(\mathbf{C})$ let $\gamma(P) = \begin{bmatrix} 0 & \mathrm{adj}(P) \\ P & 0 \end{bmatrix}$. Then

$$\sum_{k=0}^{3} p_k \gamma_k = \begin{bmatrix} 0 & \mathrm{adj}(\underline{p}) \\ \underline{p} & 0 \end{bmatrix} = \gamma(\underline{p}). \tag{100}$$

The representation λ of $H = SU(2)$ corresponding to $s = \frac{1}{2}$ is just its standard representation on $W = \mathbf{C}^2$, i.e. $\lambda(A)w = Aw$ for all $w \in W$. Consider the direct sum representation $\lambda \oplus \lambda$ on $W \oplus W = \mathbf{C}^4$ given by

$$(\lambda \oplus \lambda)(A)(w_1, w_2) = (Aw_1, Aw_2),$$

i.e.

$$(\lambda \oplus \lambda)(A) = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}.$$

Let τ be the representation of $G = SL(2, \mathbf{C})$ on $W \oplus W$ given by

$$\tau(A) = \begin{bmatrix} (A^*)^{-1} & 0 \\ 0 & A \end{bmatrix}.$$

Note that

$$\gamma(APA^*) = \tau(A)\gamma(P)\tau(A)^{-1} \tag{101}$$

for all $A, P \in SL(2, \mathbf{C})$, and

$$\tau(A) = (\lambda \oplus \lambda)(A)$$

for all $A \in SU(2)$. We next embed $\Gamma(W_\lambda)$ in $\Gamma\left((W \oplus W)_{\lambda \oplus \lambda}\right)$:

$$\begin{aligned} \Gamma(W_\lambda) &\cong \left\{\phi = (\phi_1, \phi_2) : G \to W \oplus W \ : \ \phi_i(gh) = \lambda(h^{-1})\phi_i(g) \text{ for } i = 1, 2 \ \& \ \phi_1 = \phi_2\right\} \\ &= \left\{\phi : G \to W \oplus W \ : \ \phi(gh) = \lambda(h^{-1})\phi(g) \ \& \ \gamma\left(\begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix}\right)\phi = m\phi\right\} \\ &= \left\{\phi \in \Gamma\left((W \oplus W)_{\lambda \oplus \lambda}\right) : \gamma\left(\begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix}\right)\phi = m\phi\right\} \subset \Gamma\left((W \oplus W)_{\lambda \oplus \lambda}\right). \end{aligned}$$

Using the method of Subsection 6.3 with $x_0 = \begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix}$, we observe that the isomorphism

$$\Psi : \Gamma\left((W \oplus W)_{\lambda \oplus \lambda}\right) \to C(X_m^+, W \oplus W)$$

is given by

$$\Psi\phi\left(A \begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix} A^*\right) = \tau(A)\phi(A)$$

for $A \in SL(2, \mathbf{C})$. It follows that

$$\begin{aligned}
\Psi\phi\left(A \begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix} A^*\right) &= \tau(A)\phi(A) \\
&= m^{-1}\tau(A)\gamma\left(\begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix}\right)\phi(A) \\
&= m^{-1}\tau(A)\gamma\left(\begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix}\right)\tau(A)^{-1}\tau(A)\phi(A) \qquad (102) \\
&= m^{-1}\gamma\left(A \begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix} A^*\right)\tau(A)\phi(A) \\
&= m^{-1}\gamma\left(A \begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix} A^*\right)\Psi\phi\left(A \begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix} A^*\right).
\end{aligned}$$

Writing $A \begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix} A^* = \underline{p}$ and $\Psi\phi = f$, it follows that

$$\gamma(\underline{p})f(\underline{p}) = mf(\underline{p}). \qquad (103)$$

Multiplying (103) on the left by $\gamma(\underline{p})$ we obtain

$$\left(\langle p, p\rangle - m^2\right) f(\underline{p}) = 0.$$

It follows that $f(\underline{p})d\mu_m^+$ may be viewed as a tempered distribution on $\mathbf{R}^{1,3}$. Using (103), it follows that $u = \mathcal{F}\left(f(\underline{p})d\mu_m^+\right)$ satisfies the *Dirac Equation*:

$$i\sum_{k=0}^{3} \epsilon_k \gamma_k \partial_k u = mu. \qquad (104)$$

Abbreviating $\partial\!\!\!/ = \sum_{k=0}^3 \epsilon_k \gamma_k \partial_k$, the Dirac equation reads $i\partial\!\!\!/ u = mu$. Note that $\partial\!\!\!/^2 = \Box^2$, hence again the Dirac equation implies the Klein-Gordon equation.

7 Introduction to Quantum Computing

7.1 Classical and Quantum Circuits

A classical circuit is a directed graph with n input bits $x_1, \ldots, x_n$ and m output bits $y_1, \ldots, y_m$, where in each internal vertex v there is a Boolean function, taking as inputs the bits coming from the edges directed into v, and outputing the result to the edges directed from v. It is assumed that the in-degree of each vertex is bounded, say by 2, i.e. we allow Boolean functions with at most 2 inputs. See Figure 7 for an example of a classical circuit. The complexity of the circuit is the number of gates, i.e. internal vertices.

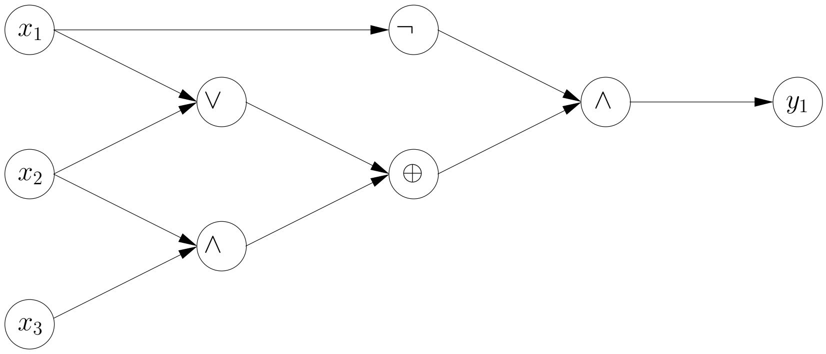


Figure 7: Classical circuit

We now define quantum circuits. Recall that a qubit is a state in $\mathbf{C}^2$. Let $e_0 = (1,0), e_1 = (0,1)$ be the standard basis of $\mathbf{C}^2$. For $\epsilon = (\epsilon_1, \ldots, \epsilon_n) \in \{0,1\}^n$ let $e_\epsilon = e_{\epsilon_1} \otimes \cdots \otimes e_{\epsilon_n}$. An *d-qubit quantum gate* is a unitary transformation in $U(\mathcal{H}^{\otimes d}) = U(2^d)$. There are of course just two classical unary gates: $\mathrm{Id}(\epsilon) = \epsilon$ and $\neg(\epsilon) = 1 - \epsilon$. On the other hand, all elements $U(2)$ are quantum unary gates. Two commonly used unary operators are the counterpart of negation $\sigma^x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, and the Hadamard gate $H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. An operator $U \in U(\mathcal{H}^{\otimes n})$ *depends on the coordinate set* $I \subset [n]$, if the projection of Ue_ϵ on the i-coordinate is e_{ϵ_i} for any $i \notin I$. A *quantum circuit* with n-qubits input is a sequence of unitary operators $(U_1, \ldots, U_m)$ in $U(\mathcal{H}^{\otimes n})$ such that each U_i depends on a bounded number, say at most 3, of the coordinates. The complexity of the circuit is the number m of operators involved.

7.2 Generalized Toffoli Gates

The *controlled not* function CNOT:$\{0,1\}^2 \to \{0,1\}^2$ is defined by

$$CNOT(x_1, x_2) = \begin{cases} (x_1, x_2) & x_1 = 0, \\ (x_1, 1 - x_2) & x_1 = 1. \end{cases}$$

The Toffoli function $CCNOT : \{0,1\}^3 \to \{0,1\}^3$ is defined by

$$CCNOT(x_1, x_2, x_3) = \begin{cases} (x_1, x_2, x_3) & x_1x_2 = 0, \\ (x_1, x_2, 1 - x_3) & x_1x_2 = 1. \end{cases}$$

Both CNOT and CCNOT are permutation on their domains, and hence can be regarded as quantum gates. More generally, for $U \in U(\mathcal{H})$, we define the *generalized Toffoli gate* $\wedge^k(U)$ to be the quantum $(k+1)$-gate whose action on the basis elements $\{e_\epsilon : \epsilon = (\epsilon_1, \ldots, \epsilon_{k+1}) \in \{0,1\}^k\}$ is given by

$$\begin{aligned} \wedge^k(U)e_\epsilon &= e_{\epsilon_1} \otimes \cdots \otimes e_{\epsilon_k} \otimes U^{\epsilon_1 \cdots \epsilon_k} e_{\epsilon_{k+1}} \\ &= \begin{cases} e_\epsilon & \epsilon_1 \cdots \epsilon_k = 0, \\ e_{\epsilon_1} \otimes \cdots \otimes e_{\epsilon_k} \otimes U e_{\epsilon_{k+1}} & \epsilon_1 \cdots \epsilon_k = 1. \end{cases} \end{aligned} \tag{105}$$

Example 7.1. *Let $X, Y \in U(\mathcal{H})$ such that*

$$XYX^{-1}Y^{-1} = i\sigma^x = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

For example:

$$X = \frac{1}{\sqrt{2}} \begin{bmatrix} -i & -1 \\ 1 & i \end{bmatrix} \quad , \quad Y = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Figure 8 depicts a realization of the Toffoli gate $\wedge^2(i\sigma^x)$ as a product of 4 unitary 2-ary gates, $\wedge(Y^{-1})$, $\wedge(X^{-1})$, $\wedge(Y)$ and $\wedge(X)$.

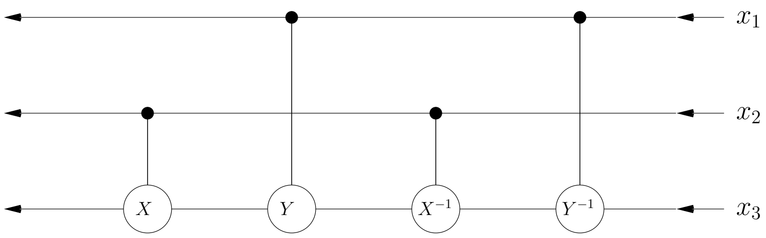


Figure 8: $\wedge^2(i\sigma^x)$ as a product of four 2-ary gates

Claim 7.2. *The generalized Toffoli gate $\wedge^k(i\sigma^x)$ can be realized as a product of $O(k^2)$ unary and 2-ary gates.*

Proof: see exercise 22.

7.3 Grover's Algorithm

Let $N = 2^n$ and let $\omega \in \{0,1\}^n$. Suppose that $f : \{0,1\}^n \to \{0,1\}$ such that $f(\omega) = 1$, and $f(\epsilon) = 0$ for $\epsilon \neq \omega$. Finding ω classically requires N queries of f.

Grover's algorithm is a quantum algorithm that determines ω with high probability by executing $\tilde{O}(\sqrt{N})$ operations. We need some preliminaries. Let $\mathcal{H} = \mathbf{C}^2$ and let $H \in U(\mathcal{H})$ be the Hadamard operator. Let $U_\omega \in U(\mathcal{H}^{\otimes n} \otimes \mathcal{H})$ be given by

$$U_\omega(e_\epsilon \otimes e_j) = e_\epsilon \otimes e_{j+f(\epsilon)}.$$

We view U_ω as the oracle for f. In the classical setting, we can ask the oracle whether $\epsilon = \omega$ and get a yes or no answer. In the quantum setting, the oracle provides us with a black box that computes the operator U_ω. We are of course not allowed to look into this black box, never the less it is crucial that the oracle constructs it with at most 2-ary gates whose number is polynomial in the size of the input. Let $\omega = (\omega_1, \ldots, \omega_n)$, and let

$$A = (\sigma^x)^{1-\omega_1} \otimes \cdots \otimes (\sigma^x)^{1-\omega_n} \otimes I.$$

Then

$$U_\omega = A \cdot \wedge^n(\sigma^x) \cdot A.$$

Note that knowing ω, the oracle can construct A as a product of n unary operators. Moreover, $\wedge^n(\sigma^x)$ is a product of $O(n^2)$ 2-ary operators. Let

$$\psi := H^{\otimes n} e_{\underline{0}} = \frac{1}{\sqrt{N}} \sum_{\epsilon \in \{0,1\}^n} e_\epsilon.$$

Let $v_\omega = e_\omega \otimes He_1$, $v_\psi = \psi \otimes He_1$, and $W = \operatorname{span}\{v_\omega, v_\psi\}$. Then

$$\begin{aligned} U_\omega v_\omega &= U_\omega(e_\omega \otimes He_1) \\ &= U_\omega\left(e_\omega \otimes \frac{1}{\sqrt{2}}(e_0 - e_1)\right) \\ &= e_\omega \otimes \frac{1}{\sqrt{2}}(e_1 - e_0) = -e_\omega \otimes He_1 = -v_\omega. \end{aligned} \tag{106}$$

Furthermore

$$\begin{aligned} U_\omega v_\psi &= U_\omega(\psi \otimes He_1) \\ &= U_\omega\left(\left(\frac{1}{\sqrt{N}} \sum_{\epsilon \in \{0,1\}^n} e_\epsilon\right) \otimes \left(\frac{1}{\sqrt{2}}(e_0 - e_1)\right)\right) \\ &= \frac{1}{\sqrt{N}} \sum_{\epsilon \in \{0,1\}^n} e_\epsilon \otimes \left(\frac{1}{\sqrt{2}}(e_{f(\epsilon)} - e_{1+f(\epsilon)})\right) \\ &= \psi \otimes He_1 - \frac{2}{\sqrt{N}} e_\omega \otimes He_1 \\ &= \left(\psi - \frac{2}{\sqrt{N}} e_\omega\right) \otimes He_1 \\ &= -\frac{2}{\sqrt{N}} v_\omega + v_\psi. \end{aligned} \tag{107}$$

Thus W is invariant under U_ω, and the matrix representing U_ω with respect to the basis $\{v_\omega, v_\psi\}$ is

$$A_\omega = \begin{bmatrix} -1 & -\frac{2}{\sqrt{N}} \\ 0 & 1 \end{bmatrix}. \tag{108}$$

Geometrically, A_ω is the reflection in the plane W in the axis spanned by

$$v := v_\omega^\perp = \frac{1}{\sqrt{N-1}} \sum_{\epsilon \neq \omega} e_\epsilon \otimes He_1 = \frac{1}{\sqrt{N-1}} \left(-v_\omega + \sqrt{N} v_\psi \right).$$

Indeed, by (108) $A_\omega v_\omega = -v_\omega$ and

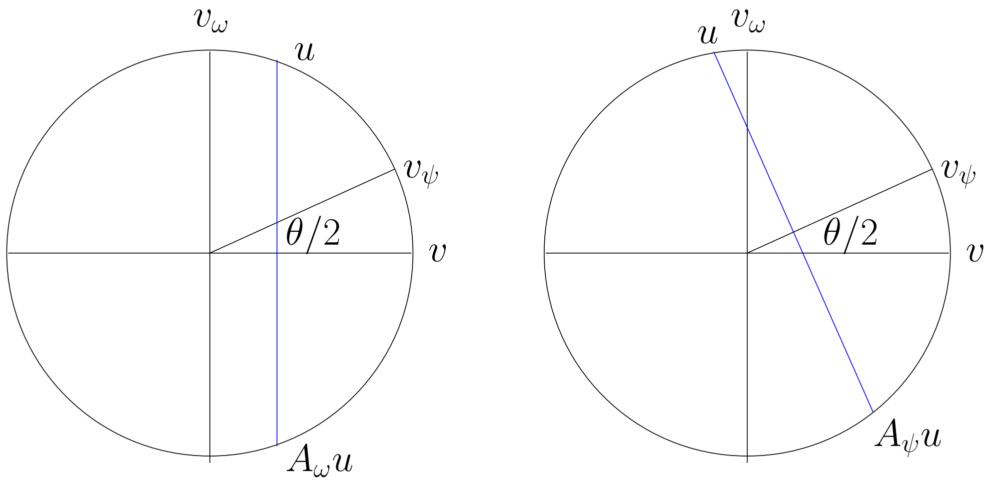


Figure 9: A_ω and A_ψ

$$\begin{aligned} A_\omega v &= \frac{1}{\sqrt{N-1}} A_\omega \left(-v_\omega + \sqrt{N} v_\psi \right) \\ &= \frac{1}{\sqrt{N-1}} \left(v_\omega + \sqrt{N} \left(-\frac{2}{\sqrt{N}} v_\omega + v_\psi \right) \right) \\ &= \frac{1}{\sqrt{N-1}} \left(-v_\omega + \sqrt{N} v_\psi \right) = v. \end{aligned} \tag{109}$$

Next, let $U_\psi \in U(\mathcal{H}^{\otimes n} \otimes \mathcal{H})$ be given by

$$U_\psi = (2\psi\psi^* - I) \otimes I.$$

Then

$$\begin{aligned} U_\psi v_\omega &= U_\psi (e_\omega \otimes He_1) \\ &= -e_\omega \otimes He_1 + \frac{2}{\sqrt{N}} \psi \otimes He_1 = -v_\omega + \frac{2}{\sqrt{N}} v_\psi \end{aligned} \tag{110}$$

and

$$U_\psi v_\psi = U_\psi(\psi \otimes He_1) = \psi \otimes He_1 = v_\psi. \tag{111}$$

Thus W is invariant under U_ψ, and the matrix representing U_ψ with respect to the basis $\{v_\omega, v_\psi\}$ is

$$A_\psi = \begin{bmatrix} -1 & 0 \\ \frac{2}{\sqrt{N}} & 1 \end{bmatrix}.$$

Geometrically, A_ψ is the reflection in the axis spanned by v_ψ. Indeed, $A_\psi v_\psi = v_\psi$ and

$$v_\psi \cdot A_\psi v_\omega = v_\psi \cdot \left(-v_\omega + \frac{2}{\sqrt{N}} v_\psi\right) = -\frac{1}{\sqrt{N}} + \frac{2}{\sqrt{N}} = \frac{1}{\sqrt{N}} = v_\psi \cdot v_\omega.$$

The *Grover operator* is $G = U_\psi U_\omega$. Thus G is the anti-clockwise rotation in the plane W with angle θ, where $\frac{\theta}{2}$ is the angle between v_ψ and v. Note that $\sin\frac{\theta}{2} = (v_\omega, v_\psi) = \frac{1}{\sqrt{N}}$, hence $\theta = 2\arcsin\frac{1}{\sqrt{N}}$. Let

$$k = \left\lfloor \frac{\pi}{2\theta} - \frac{1}{2} \right\rfloor.$$

Grover's Algorithm:

- Generate $\psi = H^{\otimes n} e_{\underline{0}} \otimes He_1$.
- Compute $G^k \psi$.
- Measure $G^k\psi$ according to the orthonormal basis $\{e_\epsilon : \epsilon \in \{0,1\}^n\}$, outputing some e_ϵ.

Proposition 7.3.

$$p := \Pr[\text{ Grover's algorithm outputs } e_\omega] \geq 1 - \frac{4}{N}.$$

Proof. $\psi = \cos\left(\frac{\theta}{2}\right) v + \sin\left(\frac{\theta}{2}\right) v_\omega$. It follows that

$$G^k\psi = \cos\left(\left(k + \frac{1}{2}\right)\theta\right) v + \sin\left(\left(k+\frac{1}{2}\right)\theta\right) v_\omega.$$

Note that

$$\frac{\pi}{2\theta} - \frac{3}{2} < k \leq \frac{\pi}{2\theta} - \frac{1}{2},$$

hence

$$\frac{\pi}{2} - \theta < \left(k + \frac{1}{2}\right)\theta \leq \frac{\pi}{2}.$$

Therefore

$$
\begin{aligned}
p &= \sin^2\left(\left(k+\frac{1}{2}\right)\theta\right) > \sin^2\left(\frac{\pi}{2}-\theta\right) \\
&= \cos^2\theta = \cos^2\left(2\arcsin\left(\frac{1}{\sqrt{N}}\right)\right) \\
&= \left(1-2\sin^2\left(\arcsin\left(\frac{1}{\sqrt{N}}\right)\right)\right)^2 \\
&= \left(1-\frac{2}{N}\right)^2 > 1-\frac{4}{N}.
\end{aligned} \tag{112}
$$

□

We next discuss the complexity of Grover's algorithm.

Proposition 7.4. *The Grover algorithm can be implement with $O(n^2\sqrt{N})$ quantum gates.*

Proof. The first step, i.e. the generation of ψ is carried out by $n+1$ application of the unary gate H. Each iteration of Grover gate $G = U_\psi U_\omega$ requires one call for the oracle operator U_ω, and an application of $U_\psi = (2\psi\psi^* - I) \otimes I$. Now

$$2\psi\psi^* - I = 2\left(H^{\otimes n}e_0\right)\left(H^{\otimes n}e_0\right)^* - I = H^{\otimes n}\left(2e_0e_0^* - I\right)H^{\otimes n}.$$

It remains to show that the operator $2e_0e_0^* - I$ can be represented as a product of $O(n^2)$ unitary operators that depend on a bounded number of coordinates. Consider the generalized Toffoli gate $T = \wedge^{n-1}(\sigma^x) \in U(\mathcal{H}^{\otimes n})$, i.e.

$$Te_\epsilon = e_{\epsilon_1} \otimes \cdots \otimes e_{\epsilon_{n-1}} \otimes e_{\epsilon_n + \epsilon_1 \cdots \epsilon_{n-1}}.$$

Define $Q, R, S \in U(\mathcal{H}^{\otimes n})$ by $Q = (\sigma^x)^{\otimes n}$, $R = I^{\otimes(n-1)} \otimes H$ and

$$S = Q \cdot R \cdot T \cdot R \cdot Q.$$

Claim 7.5.

$$S = I - 2e_0e_0^*. \tag{113}$$

Proof. It suffices to check (113) on all e_ϵ for $\epsilon = (\epsilon_1, \ldots, \epsilon_n) \in \{0,1\}^n$. Now

$$RQ(e_\epsilon) = e_{1-\epsilon_1} \otimes \cdots \otimes e_{1-\epsilon_{n-1}} \otimes He_{1-\epsilon_n}.$$

It follows that if $(\epsilon_1, \ldots, \epsilon_{n-1}) \neq (0, \ldots, 0)$ then $TRQ(e_\epsilon) = RQ(e_\epsilon)$ and hence $Se_\epsilon = QRRQe_\epsilon = e_\epsilon$. On the other hand, if $(\epsilon_1, \ldots, \epsilon_{n-1}) = \underline{0} := (0, \ldots, 0)$ then

$$
\begin{aligned}
Se_{\underline{0},\epsilon_n} &= QRT\left(RQe_{\underline{0},\epsilon_n}\right) \\
&= QRT\left(e_{\underline{1}} \otimes He_{1-\epsilon_n}\right) \\
&= (-1)^{1-\epsilon_n} QRRQe_{\underline{0},\epsilon_n} \\
&= (-1)^{1-\epsilon_n} e_{\underline{0},\epsilon_n}.
\end{aligned} \tag{114}
$$

□

Now Q and R are unary operators, and in Claim 7.2 it is shown that the Toffoli gate T is a product of $O(n^2)$ unitary operators that depend on at most 2 coordinates.

□

7.4 The Classical Fourier Transform

Let G be a finite group and let $\rho_1, \ldots, \rho_t$ denote the unitary irreducible representations of G. We view ρ_k as a homomorphism from G to $U(V_i)$, where $V_i = \mathbf{C}^{d_i}$. Let $L(G)$ denote the linear space of complex valued functions on G. The *convolution* of $f, g \in L(G)$ is given by $f * g(x) = \sum_{y \in G} f(y) g(y^{-1}x)$. The mapping $L(G) \to \mathbf{C}[G]$ given by $f \to \sum_{x \in G} f(x)x$ is an isomorphism of the algebra of functions from G to $\mathbf{C}$ with convolution, with the group algebra $\mathbf{C}[G]$. The *Fourier Transform* of a function $f \in L(G)$ is the function $\widehat{f}$ on the set of unitary representations of G, that maps a representation $\rho : G \to V_\rho$ to the endomorphism

$$\widehat{f}(\rho) = \sum_{x \in G} f(x)\rho(x) \in \mathrm{End}(V_\rho). \tag{115}$$

Claim 7.6 (Fourier Inversion Formula)**.** *For any $x \in G$*

$$f(x) = \frac{1}{|G|} \sum_{i=1}^{t} d_i tr\left(\widehat{f}(\rho_i)\rho_i(x^{-1})\right). \tag{116}$$

Proof.

$$\begin{aligned}
&\frac{1}{|G|} \sum_{i=1}^{t} d_i \mathrm{tr}\left(\widehat{f}(\rho_i)\rho_i(x^{-1})\right) \\
&= \frac{1}{|G|} \sum_{i=1}^{t} d_i \mathrm{tr}\left(\left(\sum_{y \in G} f(y)\rho_i(y)\right)\rho_i(x^{-1})\right) \\
&= \frac{1}{|G|} \sum_{i=1}^{t} d_i \sum_{y \in G} f(y)\chi_i(yx^{-1}) \\
&= \sum_{y \in G} f(y) \frac{1}{|G|} \sum_{i=1}^{t} d_i \chi_i(yx^{-1}) = f(x).
\end{aligned} \tag{117}$$

□

Claim 7.7. *The mapping $\mathcal{F} : \mathbf{C}[G] \to \prod_{k=1}^{t} \mathrm{End}(V_k)$ given by*

$$\mathcal{F}(f) = \left(\widehat{f}(\rho_1), \ldots, \widehat{f}(\rho_t)\right)$$

is an isomorphism of algebras.

Proof: The map $\mathcal{F}$ is clearly linear. Furthermore, if $f = \sum_{x\in G} f(x)x$, $g = \sum_{x\in G} g(x)x \in \mathbf{C}[G]$, then for any representation ρ of G

$$\begin{aligned}\widehat{f*g}(\rho) &= \sum_{x\in G}(f*g(x))\rho(x) \\ &= \sum_{x\in G}\sum_{y\in G} f(y)g(y^{-1}x)\rho(y)\rho(y^{-1}x) \\ &= \left(\sum_{y\in G} f(y)\rho(y)\right)\cdot\left(\sum_{z\in G} g(z)\rho(z)\right) \\ &= \widehat{f}(\rho)\cdot\widehat{g}(\rho).\end{aligned} \tag{118}$$

Claim 116 shows that $\mathcal{F}$ is injective. As

$$\dim \mathbf{C}[G] = |G| = \sum_{k=1}^{t} d_k^2 = \dim \prod_{k=1}^{t} \mathrm{End}(V_k),$$

it follows that $\mathcal{F}$ is onto. By (116), the inverse Fourier transform is given by

$$\mathcal{F}^{-1}(A_1,\ldots,A_t)(x) = \frac{1}{|G|}\sum_{i=1}^{t} d_i \mathrm{tr}\left(A_i\rho_i(x^{-1})\right). \tag{119}$$

□

Define an inner product on $\mathbf{C}[G]$ by

$$(f,g) = \sum_{x\in G} f(x)\overline{g(x)}.$$

Define an inner product on $\prod_{k=1}^{t}\mathrm{End}(V_k)$ by

$$((A_1,\ldots,A_t),(B_1,\ldots,B_t)) = \frac{1}{|G|}\sum_{k=1}^{t} d_k \mathrm{tr}(A_k B_k^*).$$

Claim 7.8 (Parseval Formula)**.** *For any $f, g \in \mathbf{C}[G]$*

$$(f,g) = (\mathcal{F}(f),\mathcal{F}(g)). \tag{120}$$

Let ρ be a representation of G on a space V, and let $V^G = \{v \in V : \rho(g)v = v \text{ for all } g \in G\}$. Let $P_{V^G} : V \to V^G$ denote the projection. The formula for projection into isotypic subspaces implies that $\sum_{x\in G}\rho(x) = |G|\cdot P_{V^G}$. Let H be a subgroup of G. The restriction of ρ to H is denoted by $\mathrm{res}_H^G\rho$. Let 1_A denote the indicator function of a set $A \subset G$. Then

$$\widehat{1_H}(\rho) = \sum_{x\in H}\rho(x) = |H|\cdot P_{V^H}. \tag{121}$$

Example 7.9 (Fourier Transform on Finite Abelian Groups)**.** *Let G be a finite abelian group. Let $\widehat{G}$ be the group of characters of G. By (115), the Fourier transform of $f \in L(G)$ is the function $\widehat{f} \in L(\widehat{G})$ given by $\widehat{f}(\chi) = \sum_{x\in G} f(x)\chi(x)$. For a subgroup $H < G$ let $H^\perp = \{\chi \in \widehat{G} : \chi(h) = 1 \text{ for all } h \in H\}$. Then by (122)*

$$\widehat{1_H}(\chi) = |H| \cdot 1_{H^\perp}(\chi). \tag{122}$$

7.5 The Hidden Subgroup Problem

Let G be a finite group and let K be a subgroup of G. Suppose there is a Hilbert space $\mathcal{H}$ and an oracle that computes a function f from G to the unit sphere of $\mathcal{H}$ such that $f(g_1) = f(g_2)$ if $g_1K = g_2K$, and $f(g_1) \perp f(g_2)$ otherwise.

The Hidden Subgroup Problem: Compute a generating set for G.

We will first consider the simplest case $G = \mathbf{F}_2^n$. A subgroup $K < G$ is a linear subspace. Let $d = \dim K$, then $\dim G/K = n - d$. Suppose $m = 2^{o(n-d)}$. If we ask a classical oracle for the value of $f(g_1), \ldots, f(g_m)$ for random elements $g_1, \ldots, g_m$, then with probability $1 - o(1)$, $g_1 + K, \ldots, g_m + K$ will be distinct and hence we will not be able to determine even a single element of K. The quantum situation is different, and in fact there is a simple polynomial time quantum algorithm that determines K. We may assume that $f : G \to \{0,1\}^{n-d}$ that satisfies $f(x) = f(y)$ iff $x - y \in K$. The quantum oracle is the unitary operator U on $(\mathbf{C}^2)^{\otimes n} \otimes (\mathbf{C}^2)^{\otimes(n-d)}$ given by

$$U(e_\epsilon \otimes e_\lambda) = e_\epsilon \otimes e_{\lambda + f(\epsilon)}.$$

Simon's Algorithm: Initialize $S = \emptyset$. Repeat $2n$ times the following steps:

- Generate $\psi = H^{\otimes n} e_{\underline{0}} \in (\mathbf{C}^2)^{\otimes n}$.
- Compute $\phi = U(\psi \otimes e_{\underline{0}})$.
- Compute $\varphi = \left(H^{\otimes n} \otimes I^{\otimes(n-d)}\right) \phi$.
- Measure the left coordinate of φ according to the standard basis $\{e_\epsilon : \epsilon \in \mathbf{F}_2^n\}$ of $(\mathbf{C}^2)^{\otimes n}$, outputing some e_ϵ.
- $S \leftarrow S \cup \{\epsilon\}$.

Proposition 7.10.

$$\Pr\left[S \text{ spans } K^\perp\right] \geq 1 - \frac{n}{2^n}.$$

Proof. Let $G = \bigcup_{i=1}^{2^{n-d}} (g_i + K)$ be the decomposition of G into cosets of K. We proceed to compute the final state of each iteration.

$$\phi = U(\psi \otimes e_{\underline{0}}) = U\left(\frac{1}{\sqrt{N}} \sum_{\epsilon \in \mathbf{F}_2^n} e_\epsilon \otimes e_{\underline{0}}\right) = \frac{1}{\sqrt{N}} \sum_{\epsilon \in \mathbf{F}_2^n} e_\epsilon \otimes e_{f(\epsilon)}. \tag{123}$$

Hence

$$\begin{aligned}
\varphi = \left(H^{\otimes n} \otimes I^{\otimes(n-d)}\right)\phi &= \frac{1}{\sqrt{N}} \sum_{\epsilon \in \mathbf{F}_2^n} \left(H^{\otimes n} e_\epsilon\right) \otimes e_{f(\epsilon)} \\
&= \frac{1}{\sqrt{N}} \sum_{\epsilon \in \mathbf{F}_2^n} \left(\frac{1}{\sqrt{N}} \sum_{\lambda \in \mathbf{F}_2^n} (-1)^{\epsilon \cdot \lambda} e_\lambda\right) \otimes e_{f(\epsilon)} \\
&= \frac{1}{\sqrt{N}} \sum_{i=1}^{2^{n-d}} \sum_{\epsilon \in K} \left(\frac{1}{\sqrt{N}} \sum_{\lambda \in \mathbf{F}_2^n} (-1)^{(\epsilon + g_i)\cdot \lambda} e_\lambda\right) \otimes e_{f(\epsilon)} \\
&= \frac{1}{2^n} \sum_{i=1}^{2^{n-d}} \left(\sum_{\lambda \in \mathbf{F}_2^n} (-1)^{g_i \cdot \lambda} e_\lambda \cdot \left(\sum_{\epsilon \in K} (-1)^{\epsilon \cdot \lambda}\right)\right) \otimes e_{f(g_i)} \\
&= \frac{2^d}{2^n} \sum_{i=1}^{2^{n-d}} \left(\sum_{\lambda \in K^\perp} (-1)^{g_i \cdot \lambda} e_\lambda\right) \otimes e_{f(g_i)} \\
&= \frac{1}{\sqrt{2^{n-d}}} \sum_{\lambda \in K^\perp} e_\lambda \otimes \left(\frac{1}{\sqrt{2^{n-d}}} \sum_{i=1}^{2^{n-d}} (-1)^{g_i \cdot \lambda} e_{f(g_i)}\right).
\end{aligned} \tag{124}$$

It follows that measuring the left coordinate of φ, we obtain a uniformly distributed random element of $K^\perp$. It follows that

$$\begin{aligned}
\Pr\left[S \text{ spans } K^\perp\right] &= \frac{(2^{2n}-1)\cdots(2^{2n}-2^{n-d-1})}{2^{2n(n-d)}} \\
&\geq \frac{(2^{2n}-2^{n-d-1})^{n-d}}{2^{2n(n-d)}} \geq \left(1 - \frac{1}{2^n}\right)^n \\
&\geq 1 - \frac{n}{2^n}.
\end{aligned} \tag{125}$$

□

We now turn to the HSP for general finite abelian groups. The approach is similar to Simon's algorithm. Let G be a finite abelian group of order N, and let K be the hidden subgroup. The oracle has a function $f : G \to G$ such that $f(g_1) = f(g_2)$ if $g_1 + K = g_2 + K$, and $f(g_1) \neq f(g_2)$ otherwise. Let $g_1, \ldots, g_N$ and let $\chi_1, \ldots, \chi_N$ be arbitrary fixed numbering of the elements of G and $\widehat{G}$, respectively. Let $\mathcal{H}$ be an N-dimensional complex Hilbert space with orthonormal basis $\{e_1, \ldots, e_N\}$. If $g = g_i$, let $\underline{g} = e_i$. If $\chi = \chi_i$, let $\underline{\chi} = e_i$. The oracle provide us with a black box unitary operator $U_K \in U(\mathcal{H} \otimes \mathcal{H})$ given by $U_K(\underline{g} \otimes \underline{h}) = \underline{g} \otimes \overline{h + f(g)}$. This operator can be realized by an efficient quantum circuit (details later). The Quantum Fourier Transform is the operator $F \in U(\mathcal{H})$ given on the basis vectors by

$$F(\underline{g}) = \frac{1}{\sqrt{N}} \sum_{\chi \in \widehat{G}} \chi(g) \underline{\chi}.$$

HSP Algorithm for abelian G : Initialize $S = \emptyset$. Repeat $r = 2 \log_2 N$ times the following steps:

- Generate $\psi = \frac{1}{\sqrt{N}} \sum_{g \in G} \underline{g}$.
- Compute $\phi = U_K (\psi \otimes \underline{0})$.
- Compute $\varphi = (F \otimes I^n) \phi$.
- Measure the left coordinate of φ according to the standard basis $\{\underline{\chi} : \chi \in \widehat{G}\}$ of $\mathcal{H}$, outputing some $\underline{\chi}$.
- $S \leftarrow S \cup \{\chi\}$.

Proposition 7.11.

$$\Pr \left[S \text{ generates } K^{\perp} \right] \geq 1 - \frac{1}{N}.$$

Proof. Let $m = \frac{|G|}{|K|}$ and let $G = \bigcup_{i=1}^{m} (g_i + K)$ be the decomposition of G into posets of K. We compute the final state of each iteration.

$$\phi = U_K (\psi \otimes \underline{0}) = \frac{1}{\sqrt{N}} \sum_{g \in G} \underline{g} \otimes \underline{f(g)}.$$

Hence

$$
\begin{aligned}
\varphi &= (F \otimes I^n)\,\phi \\
&= (F \otimes I^n)\left(\frac{1}{\sqrt{N}}\sum_{g\in G}\underline{g}\otimes\underline{f(g)}\right) \\
&= \frac{1}{\sqrt{N}}\sum_{g\in G}\left(\frac{1}{\sqrt{N}}\sum_{\chi\in\widehat{G}}\chi(g)\underline{\chi}\right)\otimes\underline{f(g)} \\
&= \frac{1}{N}\sum_{\chi\in\widehat{G}}\sum_{g\in G}\Big(\chi(g)\underline{\chi}\otimes\underline{f(g)}\Big) \\
&= \frac{1}{N}\sum_{i=1}^{m}\sum_{h\in K}\sum_{\chi\in\widehat{G}}\Big(\chi(g_i+h)\underline{\chi}\otimes\underline{f(g_i)}\Big) \qquad (126)\\
&= \frac{1}{N}\sum_{\chi\in\widehat{G}}\left(\sum_{h\in K}\chi(h)\right)\underline{\chi}\otimes\left(\sum_{i=1}^{m}\chi(g_i)\underline{f(g_i)}\right) \\
&= \frac{|K|}{N}\sum_{\chi\in K^{\perp}}\underline{\chi}\otimes\left(\sum_{i=1}^{m}\chi(g_i)\underline{f(g_i)}\right) \\
&= \frac{1}{\sqrt{m}}\sum_{\chi\in K^{\perp}}\underline{\chi}\otimes\left(\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\chi(g_i)\underline{f(g_i)}\right).
\end{aligned}
$$

It follows that on measuring the left coordinate of φ, we obtain a uniformly distributed random element of $K^{\perp}$. Therefore

$$\Pr\left[S \text{ generates } K^{\perp}\right] \geq 1 - \frac{|K^{\perp}|}{2^r} = 1 - \frac{1}{N}.$$

□

7.6 Shor's Factoring Algorithm

The efficient quantum algorithm for factoring, due to Shor, depends on a certain special case of the abelian HSP that we now describe. Let N be an positive integer, and let $\mathbf{Z}_N^*$ denote the multiplicative group of invertible elements in the ring $\mathbf{Z}_N$. For an element $a \in \mathbf{Z}_N^*$, let $\mathrm{ord}_N(a)$ denote the order of a in $\mathbf{Z}_N^*$, i.e. the minimal $r \geq 1$ such that $a^r \equiv 1 (\bmod N)$. Let

$$A_N = \left\{a \in \mathbf{Z}_N^* : \mathrm{ord}_N(a) = r \text{ is even } \ \& \ a^{\frac{r}{2}} \not\equiv -1 (\bmod N)\right\}.$$

Claim 7.12. *Suppose $N = pq$. Then:*

(i) $\Pr[A_N] \geq \frac{1}{2}$.

(ii) If $a \in A_N$ then

$$D = \left\{\gcd(a^{\frac{r}{2}} - 1, N), \gcd(a^{\frac{r}{2}} + 1, N)\right\} = \{p, q\}.$$

Proof. (i) Later.
(ii) $a \in A_N$ implies that $B \subset \{1, p, q\}$. However, if say $p \notin B$ then $a^r - 1 = \left(a^{\frac{r}{2}} - 1\right)\left(a^{\frac{r}{2}} + 1\right)$ is coprime to p, a contradiction.

□

Claim 7.12 facilitates a simple probabilistic factoring algorithm for $N = pq$, provided that we can efficiently find the order of an element:

- Choose a random element $1 \le a \le N - 1$.
- Compute $d = \gcd(a, N)$. If $1 < d$ then $d = p$ or $d = q$ and we halt. Otherwise:
- Compute $\mathrm{ord}_N(a) = r$.
- If r is odd or $a^{\frac{r}{2}} \equiv -1 (\mathrm{mod}\, N)$, go to the first step. Otherwise:
- $$\left\{\gcd(a^{\frac{r}{2}} - 1, N), \gcd(a^{\frac{r}{2}} + 1, N)\right\} = \{p, q\}.$$

Repeating the basic iteration s times, the algorithm succeeds with probability at least $1 - 2^{-s}$ in factoring N.

We now show that finding the order of an element a modulo N can be achieved by a variation on the abelian HSP. We first recall the following classical fact. Let k be fixed. For $n \ge 1$ let

$$D_k(n) = |\{(a_1, \ldots, a_k) \in [n]^k : \gcd(a_1, \ldots, a_k) = 1\}|.$$

Claim 7.13.

$$\liminf_{n \to \infty} \frac{D_k(n)}{n^k} = \zeta(k)^{-1} = \left(\sum_{j=1}^{\infty} \frac{1}{j^k}\right)^{-1} > 1 - 2^{-(k-1)}.$$

Quantum algorithm for finding $r = \mathrm{ord}_N(a)$.

Choose an integer M such that $M \approx N^2$. Let $G = \mathbf{Z}_M$ and let $K = r\mathbf{Z}_M$ be the hidden subgroup. Let $f : \mathbf{Z}_M \to \mathbf{Z}_M$ be given by $f(x) = a^x (\mathrm{mod}\, N)$. For the sequel we assume that M is divisible by r. We of course cannot guarantee this a priori, but it turns out that choosing $M \approx N^2$ gives a sufficiently good approximation for r. Now run the hidden subgroup algorithm for $\mathbf{Z}_M$ and the above function f. The algorithm outputs a uniformly chosen set $\{a_1, \ldots, a_k\}$ in the subgroup $r\mathbf{Z}_M$. By Claim 7.13, the probability that $\gcd(a_1, \ldots, a_k) = r$ is $\ge 1 - 2^{-(k-1)}$.