Appendices

9.10 Non-examinable group and representation theory details/proofs

9.10.1 Properties of functions

Consider two sets, X and Y. A *function* (or *map*) f from X to Y is defined such that, for each element x belonging to X (denoted as $x \in X$), there exists a unique element y in Y associated with x. We represent this element as $y = f(x)$ and call it the image of x under the function f. We write it as:

$$f: \; X \to Y \quad , \quad x \mapsto y = f(x) \, . \tag{9.95}$$

The set X is called the *domain* of f, and Y is its *image*. The set of elements in Y, which are images under f of elements in X, is called the image of X under f and is denoted as $f(X)$. In general, $f(X)$ is a subset of Y (we write $f(X) \subset Y$) and is not necessarily identical to Y.

The function f is *injective* if:

$$f(x) = f(x') \;\; \Rightarrow \;\; x = x' \, . \tag{9.96}$$

For an injective function, two elements of X cannot have the same image in Y. A function is *surjective* if $f(X) = Y$. For a surjective function, every element of Y is the image of at least one element of X. A function that is both injective and surjective is called *bijective*.

Let f be a function from X to Y and g be a function from Y to Z. The *composition* or *product* of these two functions $h : X \to Z$ is defined as:

$$h(x) = g\left(f(x)\right) \, . \tag{9.97}$$

The function h acts from X to Z and is denoted as:

$$h = g * f \tag{9.98}$$

or simply gf when there is no possibility of confusion with other operations. It should be noted that $f * g$ is not necessarily well-defined, and when it exists, it is not necessarily equal to $g * f$. For example, consider real-valued functions $f(x) = x^2$ and $g(y) = \mathrm{e}^y$. We have:

$$(g * f)(x) = g\left(x^2\right) = \mathrm{e}^{x^2} \tag{9.99}$$

and

$$(f * g)(x) = f\left(\mathrm{e}^x\right) = \mathrm{e}^{2x} \, . \tag{9.100}$$

The composition of functions is associative, meaning that if u, v, and w are functions from X to Y, Y to Z, and Z to W, respectively, then:

$$(w * (v * u))(x) = ((w * v) * u)(x) \, . \tag{9.101}$$

For each $x \in X$, both sides of this equation correspond to the element:

$$w(v(u(x))) \tag{9.102}$$

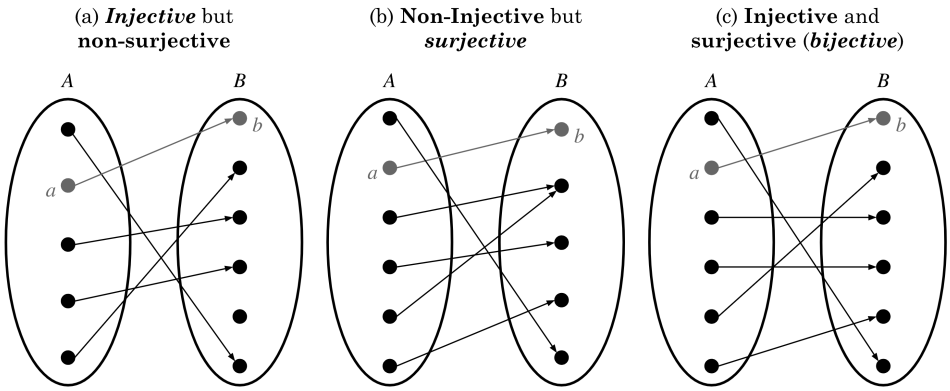

Figure 9.10: Diagram of injective, surjective and bijective functions: (Wiki page on functions).)

in W. Therefore, we can write:

$$(w * (v * u))(x) = ((w * v) * u)(x) = w * v * u\,. \tag{9.103}$$

If $f : X \to Y$ is a bijective application, then for each element y in Y, there is a unique element x in X such that $f(x) = y$, and, naturally, each element x has an image in Y. Therefore, we can define a bijective application $Y \to X$, $y \mapsto x$ such that $y = f(x)$. This application is called the *inverse* of f and is denoted by f^{-1}.

Often, we consider applications from a set X to itself. An example is given by real (complex) functions of a real (complex) variable. We define the identity application as:

$$e : \; X \to X \quad , \quad x \mapsto e(x) = x\,. \tag{9.104}$$

This application is clearly bijective. If $f : X \to Y$ is a bijective application, f^{-1} exists, and we have:

$$(f^{-1} * f)(x) = x \tag{9.105}$$

for each x. Therefore, we write:

$$f^{-1} * f = e_X \tag{9.106}$$

where we denote the identity application of X by e_X. Note that we also have:

$$f * f^{-1} = e_Y \tag{9.107}$$

Theorem. Let X and Y be two sets containing the same finite number n of elements[23]. The following three statements are equivalent:

(i) $f : X \to Y$ is surjective,

(ii) $f : X \to Y$ is injective,

(iii) $f : X \to Y$ is bijective.

Proof:

(i) $\Rightarrow f(X) = Y$. Thus, $f(X)$ is composed of n elements, which implies (ii).

[23]Note that this theorem is not valid for two sets with different numbers of elements.

(ii) $\Rightarrow$ $f(X)$ is composed of n elements. It follows that $f(X) = Y$, which can be reduced to property (i).

Since (i) and (ii) are each a consequence of the other, (iii) is also true, and the theorem is thus proved.

9.10.2 Reordering Theorem

A useful theorem in finite group theory is the reordering theorem:

Theorem 9.10.1 (Reordering theorem)**.** *Let G be a finite group and m one of its elements. The ensembles mG and Gm are a re-order of G.*

First lets check that this is true for an example we've just looked at - the rectangle symmetry group R_2. Suppose we take the set $G = \{e, a, b, c\}$ and multiply each element by the element $m = a$ to give $mG = \{a, a^2, ab, ac\}$. Then from the Cayley table of the group in Eq. 9.15 we get $mG = \{a, e, c, b\}$. This is just the original group reshuffled. As the group is Abelian the same applies for Gm.

This is the type of theorem where I almost feel that just staring and convincing yourself of it is the best way forward- as groups are closed multiplying through by a group element gives another group element and the requirement for each element have an inverse means that you always get a different element on multiplying through. Alternatively, one can more formally argue the following:

Demo. The map $x \to mx$ is surjective (i.e. all elements have an antecedent). Indeed for any $y \in G$, $m^{-1}y \in G$ (group property) and $m(m^{-1}y) = y$. The map $x \to mx$ is also injective (it maps distinct elements to distinct elements). For any $x \neq x'$, $mx \in G$ is different from mx'. Indeed, if $mx = mx'$ then $m^{-1}(mx) = m^{-1}mx'$ and $x = x'$. Hence the map is bijective and therefore a reordering. The proofs works in a similar way for the map $x \to xm$. (For more on the properties of functions see Appendix 9.10.1) □

9.10.3 Proof of Lagrange Theorem, Right and Left cosets

Let G be a group and H one of its proper subgroups. We can define an equivalence relation —different from the last one— between the elements of G as follows: if $x, y \in G$ and $x^{-1}y \in H$ then x and y are equivalent and we write $x \sim y$.

This is indeed an equivalence relation:

- $a^{-1}a = e \; \forall a \in G$, and $e \in H$, so that $a \sim a$.
- if $a \sim b$ then $a^{-1}b \in H$. The inverse of $a^{-1}b$ is $b^{-1}a$ and since H is a group, $b^{-1}a \in H$, so that $b \sim a$.
- if $a \sim b$ and $b \sim c$, then $a^{-1}b$ and $b^{-1}c$ are both in H, thus so is the their product $a^{-1}bb^{-1}c = a^{-1}c$.

This equivalence relation therefore makes it possible to divide the elements of G into disjoint classes. If $x^{-1}y \in H$, then y is equal to an element of H multiplied on the left by x. We indicate the set thus constructed by the symbol

$$C_x = xH \tag{9.108}$$

which the call the *left co-set associated to* x.

The map $H \to xH$ is one-to-one (bijective). Indeed, each element $z \in xH$ is the image of $x^{-1}z \in H$ so that the map is surjective. But the map is also injective since for $y, y' \in H$, we have $xy = xy' \Rightarrow y = y'$.

We could also define a second equivalence relation $x \sim y$ if $yx^{-1} \in H$ and which this case, we can define the concept of right co-set Hx in the same way as before.

These concepts are very useful, and allows in particular to prove Lagrange Theorem:

Demo. Consider the co-sets on the left of H. They are all disjoint or identical (since they are equivalence classes). If there are n distinct left co-sets, their union is G. So, if we denote by g and h the orders of G and H respectively, then $g = nh$ and the theorem is proved. □

Let us give an example for the following order 4 group

$$G = \begin{array}{c|cccc} * & e & a & b & c \\ \hline e & e & a & b & c \\ a & a & e & c & b \\ b & b & c & e & a \\ c & c & b & a & e \end{array} \tag{9.109}$$

that has the subgroup $H = \{e, a\}$:

$$H = \begin{array}{c|cc} * & e & a \\ \hline e & e & a \\ a & a & e \end{array} \tag{9.110}$$

We can now construct the left co-sets:

$$C_e = eH = \{e, a\} \tag{9.111}$$
$$C_a = aH = \{a, e\} = C_e \tag{9.112}$$
$$C_b = bH = \{b, c\} \tag{9.113}$$
$$C_c = cH = \{c, b\} = C_b \tag{9.114}$$

And we see indeed that we have *two* left co-set of order 2.

9.10.4 Composition of Conjugacy classes theorem statement and proof

Theorem 9.10.2 ("Composition" of conjugacy classes)**.** *Let G be a group, and C_x and C_y two of its conjugacy classes. Then we have*

$$C_\nu * C_\mu = \sum_\lambda n_{\mu\nu\lambda} C_\lambda \tag{9.115}$$

with $n_{\mu\nu\lambda}$ integer. Here the multiplication $C_\nu * C_\mu$ is defined as the entire set $[xy]$ for all $x \in C_\nu$ and $y \in C_\mu$. Additionally, $n_{\nu\mu\lambda} = n_{\mu\nu\lambda}$ and $n_{1\nu\lambda} = n_{\nu 1\lambda} = \delta_{\nu,\lambda}$.

To prove this, let us first prove a variant of the reordering theorem:

Theorem 9.10.3 (Reordering theorem within conjugacy classes)**.** *Let G be a group, m one of its elements, and C one of the conjugate classes. Then the application $C \to m^{-1}Cm$ is bijective into itself: The ensemble $m^{-1}Cm$ is thus a re-ordering of C.*

Demo. First notice that this is a map into itself since for any $y \in C$, $m^{-1}ym \in C$ (conjugacy class property). Second, the map is surjective. Indeed, for any $y \in C$, it exists $x = mym^{-1} \in G$ such that $y = m^{-1}xm$. By definition, x is thus also in C and therefore for all $y \in C$ there is an antecedent in C. Third, the map $x \to mx$ is injective (it maps distinct elements to distinct elements). For any x, x', we have $m^{-1}xm = m^{-1}x'm$ implies that $mm^{-1}xmm^{-1} = mm^{-1}x'mm^{-1}$ so that $x = x'$. □

We shall soon prove that $n_{\nu\mu\lambda}$ is indeed an integer. But first, let us note indeed that $n_{\nu\mu\lambda} = n_{\mu\nu\lambda}$, because the two sets $C_\nu * C_\mu$ and $C_\mu * C_\nu$ are identical. Indeed

$$C_\nu * C_\mu = [uv] = \left[uv(u^{-1}u)\right] = \left[u(vu^{-1}u)\right] = \left[uvu^{-1}u\right] = \left[(uvu^{-1})u\right] = C_\mu * C_\nu$$

since u represents all the element of C_ν, and since, from the previous theorem, (uvu^{-1}) represent a re-ordering of the all the element of C_μ as v changes. Additionally, we also see that, denoting the class that contains e are C_1, that $C_1 * C_\nu = C_\nu$ so that $n_{1\nu\lambda} = n_{\nu 1\lambda} = \delta_{\nu,\lambda}$.

Let us now prove that $n_{\nu\mu\lambda}$ is an integer. First we prove the following lemma:

Lemme 9.10.4. *A necessary and sufficient condition for a set $[R]$ to be composed uniquely of a set of entire classes of a group G is that*

$$\forall u \in G\,,\ u^{-1}[R]u = [R]$$

Demo. The condition is necessary because, if indeed $[R]$ is composed of entire sets, then in each of these sets S, $u^{-1}[S]u$ is itself the set S by the reordering theorem.

To see that the condition is sufficient, let us proceed by contradiction and write

$$[R] = [R'] + [R'']$$

where $[R']$ is the largest subset of $[R]$ made of entire classes, and the reminder $[R'']$ thus must contain elements that are not an entire class. Since $[R']$ satisfy $u^{-1}[R']u = [R']$ then

$$u^{-1}[R'']u = [R'']\,.$$

e cannot be in $[R'']$ since it is, itself, a class. Let us suppose $[R'']$ is not empty, and $x \in [R'']$. Then it must exists $y \in G$, conjugated to x, which is *not* in $[R'']$. Since y is conjugated to x we have $u^{-1}xu = y$ for some $u \in G$. But then since $u^{-1}[R'']u = [R'']$ for all u, y must be in $[R'']$. We have thus reach a contradiction, and $[R'']$ is empty. □

Now we can proceed. Let H be a finite group of order h and conjugacy classes $C_1 = \{e\}$, $C_2, \ldots, C_\mu, \ldots, C_{N_C}$ its classes. We shall denote by n_μ the number of elements in the class C_μ and by N_C the total number of classes. We have, of course

$$\sum_{\mu=1}^{N_c} n_\mu = h \tag{9.116}$$

Let C_μ and C_ν be two classes of H, and consider the product

$$C_\mu * C_\nu = [uv] \tag{9.117}$$

where u and v are elements of C_μ and C_ν. Then for each $x \in H$, we have

$$x^{-1}C_\mu * C_\nu x = [x^{-1}uvx] = [x^{-1}u(xx^{-1})vx] = [(x^{-1}ux)(x^{-1}vx)] \tag{9.118}$$

Using the theorem of rearrangement, we see that $[(x^{-1}ux)(x^{-1}vx)]$ is just a reordering of $[uv]$ so that

$$x^{-1}C_\mu * C_\nu x = C_\mu * C_\nu \tag{9.119}$$

Applying lemma 9.10.4 then prove theorem 9.10.2.

9.10.5 Proof of Schur's lemma

Let us prove Schur's lemma. We are going to need the definition of "kernel" and "image" of an operator.

Definition 9.10.5 (Kernel of an operator)**.** The kernel Ker$A$ of an operator $A : V_1 \to V_2$ is the set of vector $\boldsymbol{v}_1 \in V_1$ such that $Av_1 = 0$.

Definition 9.10.6 (Image of an operator)**.** The image ImA of an operator $A : V_1 \to V_2$ is the set of vector $\boldsymbol{v}_2 \in V_2$ for which $\exists v_1 \in V_1$ such that $v_2 = Av_1$.

Theorem 9.10.7 (Rank-Nullity theorem)**.** *For any operator* $A : V_1 \to V_2$*, define* $\text{Rank}(A) = \dim[\text{Im}(A)]$ *and* $\text{Nullity}(A) = \dim[\text{Ker}(A)]$*, then* $\dim[V_1] = \text{Nullity}(A) + \text{Rank}(A)$.

9.10.6 Proof of lemma 1

Demo. For all $g \in G$ we have:

- $\forall \boldsymbol{v}_1 \in \text{Ker}A$ we have $A(R_1(g)\boldsymbol{v}_1) = R_2(g)A\boldsymbol{v}_1 = 0$. This means that the vector $R_1(g)\boldsymbol{v}_1$ is also in the kernel of A. In other words a vector in $W = \ker A$ stays in W upon transformation by $R_1(g), \forall g$: W is thus a stable sub-space of $R_1(g)$.
- From a similar reasoning, we can deduce that the image $W' = \text{Im}A$ is also a stable subspace for $R_2(g)$. Indeed, this requires implies that if a vector can be written as $\boldsymbol{v}_2 = A\boldsymbol{v}_1$, then $R_2(g)\boldsymbol{v}_2$ can also be written as $A\boldsymbol{v}_1'$. This is the case since $R_2(g)\boldsymbol{v}_2 = R_2(g)A\boldsymbol{v}_1 = AR_1(g)\boldsymbol{v}_1 = A\boldsymbol{v}'$.

We thus conclude that $W = \text{Ker}A$ is a stable subspace $R_1(g)$ and that $W' = A$ is a stable subspace of $R_2(g)$. However, by assumption, both representations are irreducible, so the only subspaces are either 0 or the entire space. We thus have either:

- Ker$A = 0$, in which case the image is not empty, so that im$A = V_2$. But this implies that the transformation A is invertible, but then $A^{-1}R_2(g)A = R_1(g)\forall g$, and R_2 and R_1 are equivalent, which contradicts the hypothesis.
- Ker$A = V_1$, in which case $A = 0$ (and the image is empty:im$A = 0$).

□

9.10.7 Proof of lemma 2

In this case, we have a map between either the same, or between equivalent representations. Additionally, $V_1 = V_2 = V$, and A is a square matrix. If the representation are equivalent, we can always rotate the space so that they are indeed identical.

Let us consider then that $R_1(g) = R_2(g) = R(g)\forall g$.

Demo. By the fundamental theorem of algebra, it exists an eigenvalue $\lambda \in \mathbb{C}$ such that $\det(A - \lambda I) = 0$. Consider then the equation

$$(A - \lambda I)R(g) = R(g)(A - \lambda I)\,. \tag{9.120}$$

so that if $v \in \mathrm{Ker}(A-\lambda I)$ then $R(g)v$ also in $\mathrm{Ker}(A-\lambda I)$. $W = \mathrm{Ker}(A-\lambda I)$ is thus a stable subspace of transformation by $R(g) \forall g$. Given $R(g)$ is irreducible, either $W = 0$ or $W = V$. W cannot be zero, because at least the eigenvector of A corresponding to λ is in W! Therefore $W = V$.

We this have $\mathrm{Ker}(A-\lambda I) = V$, so that $(A-\lambda I) = 0$ and therefore $A = \lambda I$. □

9.10.8 Proof of grand orthogonality Lemma

Demo. Consider any matrix X and the matrix M defined as

$$M = \sum_{g \in G} R_1(g^{-1}) X R_2(g) \tag{9.121}$$

Then we have, for any $y \in G$

$$\begin{aligned} MR_2(y) &= \sum_{g \in G} R_1(g^{-1}) X R_2(g) R_2(y) \\ &= \sum_{g \in G} R_1(y) R_1(y^{-1}) R_1(g^{-1}) X R_2(g) R_2(y) \\ &= R_1(y) \sum_{g \in G} R_1(y^{-1}) R_1(g^{-1}) X R_2(g) R_2(y) \\ &= R_1(y) \sum_{g \in G} R_1(y^{-1} g^{-1}) X R_2(gy) \\ &= R_1(y) \sum_{g \in G} R_1((gy)^{-1}) X R_2(gy) \\ &= R_1(y) \sum_{h \in G} R_1(h^{-1}) X R_2(h) = R_1(y) M \end{aligned}$$

We can thus use Schur's lemmas on M. Since R_1 and R_2 are not equivalent we have $M = 0$ so that

$$\sum_{g \in G} \sum_{jl} [R_1(g^{-1})]_{kj} X_{jl} [R_2(g)]_{lm} = 0 \tag{9.122}$$

but $R_1(g^{-1})$ is $R_1(g)^\dagger$ so that

$$\begin{aligned} \sum_{g \in G} \sum_{jl} [R_1(g)]^\dagger_{kj} X_{jl} [R_2(g)]_{lm} &= 0 \\ \sum_{g \in G} \sum_{jl} [R_1(g)]^*_{jk} X_{jl} [R_2(g)]_{lm} &= 0 \end{aligned} \tag{9.123}$$

Using $X_{jl} = 0$ except for one pair jl for which $X_{jl} = 1$ leads to eq.(9.54).

We now turn to eq.(9.56). If we construct the matrix M using the same representation, we get again $MR(x) = R(x)M$ and by the second Schur lemma:

$$\sum_{g \in G} R(g^{-1}) X R(g) = c(X) I \tag{9.124}$$

which, in full matrix notation, means

$$\sum_{g \in G} \sum_{jl} [R(g)]^*_{jk} X_{jl} [R(g)]_{lm} = c(X) \delta_{km}$$

We just need to compute the constant. Let us work on the diagonal, when $k = m$, and sum over k so that we have

$$\begin{aligned}
\sum_{g\in G}\sum_{jlk}[R(g^{-1})]_{kj}X_{jl}[R(g)]_{lk} &= n_a c(X)\\
\sum_{g\in G}\sum_{jl}X_{jl}\sum_{k}[R(g^{-1})]_{kj}[R(g)]_{lk} &= n_a c(X)\\
\sum_{g\in G}\sum_{jl}X_{jl}[R(g)R(g^{-1})]_{lj} &= n_a c(X)\\
\sum_{g\in G}\sum_{jl}X_{jl}I_{lj} &= n_a c(X)\\
\sum_{g\in G}\mathrm{Tr}X &= n_a c(X)\\
c(X) &= \frac{N}{n_A}\mathrm{Tr}X
\end{aligned}$$

Using again $X_{jl} = 0$ except for one pair jl for which $X_{jl} = 1$ leads to eq.(9.56). □

9.10.9 Proof of Burnside Lemma

We can now prove Burnside lemma. Consider the regular representation (which we introduced in the previous chapter) that is obtained using $N \times N$ matrices for a finite group of order N. Then we have a amazing fact: Any irreducible representation R of G appears in the regular representation $dim(R)$ times:

Theorem 9.10.8 (Regular representation decomposition)**.** *Consider the regular representation of a group. Then we have the following decomposition in irrep*

$$R^r(g) = \oplus_{a,x}R_{a,x}(g) = \oplus_a R_a R_a(g) \tag{9.125}$$

where R_a is the dimension of the representation a.

Demo. We simply apply

$$b_a = \frac{1}{N}\sum_\mu n_\mu \chi_a^*(C_\mu)\chi^r(C_\mu) \tag{9.126}$$

and using the fact that for the regular representation all characters are zero except for the one corresponding to e, we find

$$b_a = \frac{1}{N}\chi_a^*(C_e) = \frac{R_a}{N}N = R_a \tag{9.127}$$

□

This finally allows to prove Burnside's lemma, by simply counting the dimensions:

Lemme 9.10.9 (Burnside lemma)**.**

$$\sum_{i=1}^{N_r} d_i^2 = N \tag{9.128}$$

9.10.10 Representations of Lie Groups and Lie Algebras

The following theorems hold on the relationship between representations of Lie groups and Lie algebras [24]:

[24] These statements are taken from Representation Theory for Geometric Quantum Machine Learning - see there for further discussion.

Theorem 9.10.10 ((Lie group reps induce Lie algebra reps)). *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. If R is a representation of G on V, then there exists a unique representation r of $\mathfrak{g}$ on V given by*

$$r(J) = \frac{d}{dt}(R(e^{tX}))\Big|_{t=0}, \text{ for all } X \in \mathfrak{g}. \tag{9.129}$$

We call r the representation of $\mathfrak{g}$ induced by R.

Theorem 9.10.11 ((Lie algebra reps lift to simple Lie group representations)). *Let G be a simply connected matrix Lie group, and let r be a representation of the corresponding Lie algebra on V. Then there is a unique representation R of G with the property*

$$R(e^X) = e^{r(X)}, \text{ for all } X \in \mathfrak{g}. \tag{9.130}$$

Theorem 9.10.12 ((Lie algebra reps locally lift to Lie group reps).). *Let G be a matrix Lie group, and let r be a representation of the corresponding Lie algebra on V. Then using Theorem 1, we can always locally define a representation R on G by the mapping*

$$R(g) = e^{r(X)}, \text{ defined for all } g = e^X \text{ nearby } I. \tag{9.131}$$

Here, by "nearby" we mean "wherever the exponential map is a diffeomorphism". Indeed, in this region, all g can be written as $g = e^X$.

9.10.11 Projectors

In general, Theorem 9.9.3 allows us to compute how many times any irrep appears in any representation we are trying to reduce. But we have yet to establish a general strategy for finding the *basis* in which the representation is block diagonalized to the sum of irreps. We address that in this section.

9.10.12 Basis notation

Let us start by introducing *basis notation*. As we saw in the previous section, we can take a representation and write it as a direct sum of irreducible representations. That is, we can write $R(g) = \oplus_{a,x} R(a, x(g)$, or equivalently

$$R(g) = \begin{pmatrix} R_1(g) & 0 & 0 & 0 & \dots \\ 0 & R_1(g) & 0 & 0 & \dots \\ 0 & 0 & R_2(g) & 0 & \dots \\ 0 & 0 & 0 & R_2(g) & \dots \\ \dots & & & & \end{pmatrix} = \begin{pmatrix} R_{1,1}(g) & 0 & 0 & 0 & \dots \\ 0 & R_{1,2}(g) & 0 & 0 & \dots \\ 0 & 0 & R_{2,1}(g) & 0 & \dots \\ 0 & 0 & 0 & R_{2,2}(g) & \dots \\ \dots & & & & \end{pmatrix}$$

We are going to denote the element in this basis using 3 indices as : $\{|a, j, x\rangle\}$. That is, we can write

$$\langle a, j, x| R(g) |b, k, y\rangle = \delta_{a,b}\delta_{x,y} [R_{a,x}(g)]_{jk} \tag{9.132}$$

Here $[R_{a,x}(g)]_{jk}$ is just the j, k element of the matrix for the representation $R_a(g)$. Here:

- $a = 1, 2, 3, \dots$ denote type of representation, i.e. indicate each of the non-equivalent representations $R_1, R_2, R_3, \dots$. At this point: $R_a(g)$ acts in a subspace $\mathcal{H}_a$.
- The same representation can be used multiple times, as we have seen in the previous example. The x index denotes which of these equivalent representation we consider.
- Finally; $\{|a, j, x\rangle\}$ with $j = 1, 2, 3, \dots$ is used to represent a basis within the x_{th} copy of subspace $\mathcal{H}_a$.

9.10.13 How to construct projectors

The question we address in this section is how to construct projectors onto $|a,j,x\rangle$. To do so, we start from

$$R(g) = \oplus_{a,x} R_{a,x}(g) \tag{9.133}$$

and

$$R_{a,x} = \sum_{lm} \langle a,l,x|\, R(g)\, |a,m,x\rangle |l\rangle\langle m| = \sum_{lm} [R_a(g)]_{lm} |l\rangle\langle m| \,. \tag{9.134}$$

so applying the representation to the vector $|a,j,x\rangle$ gives:

$$R(g)\,|a,j,x\rangle = \sum_k [R(a(g)]_{kj}\,|a,k,x\rangle \ . \tag{9.135}$$

Now we multiply by $[R_b(g)]^*_{k'j'}$ and sum over the group elements to give:

$$\begin{aligned} \sum_g [R_b(g)]^*_{k'j'} R(g)\,|a,j,x\rangle &= \sum_k \sum_g [R_b(g)]^*_{k'j'} [R(a(g)]_{kj}\,|a,k,x\rangle \\ &= \sum_k \sum_g [R_b(g)^\dagger]_{j'k'} [R(a(g)]_{kj}\,|a,k,x\rangle \ . \end{aligned} \tag{9.136}$$

and now, using the grand orthogonality theorem, one finds

$$\begin{aligned} \sum_g [R_b(g)]^*_{k'j'} R(g)\,|a,j,x\rangle &= \sum_k \frac{N}{n_a} \delta_{jj'}\delta_{kk'}\delta_{ab}\,|a,k,x\rangle \\ &= \frac{N}{n_a}\delta_{ab}\delta_{jj'}\,|a,k',x\rangle \ . \end{aligned} \tag{9.137}$$

Thus, we see that we can define

$$\hat{\Pi}^b_{kj} = \frac{n_a}{N} \sum_g [R_b(g)]^*_{kj} R(g) \tag{9.138}$$

such that we have

$$\hat{\Pi}^b_{kj}\,|a,j,x\rangle = \frac{n_a}{N} \sum_g [R_b(g)]^*_{kj} R(g)\,|a,j,x\rangle = \delta_{ab}\,|a,k,x\rangle \tag{9.139}$$

That is this operator satisfies:

$$\hat{\Pi}^a_{kj}\,|a,j,x\rangle = |a,k,x\rangle \tag{9.140}$$

$$\hat{\Pi}^a_{kj}\,|b,j',x\rangle = 0 \quad \text{otherwise} \tag{9.141}$$

so that if we know one vector of the basis, then we can find all the other ones! And Π^a_{kk} can be used to find that first vector.

Now, let us take the trace. We find

$$\hat{P}_a = \sum_j \hat{\Pi}^a_{jj} = \sum_g \frac{n_a}{N} \sum_j [R_b(g)]^*_{jj} R(g) = \frac{n_a}{N} \sum_g \chi^*_a(g) R(g) \tag{9.142}$$

This is a projector on the basis of the representation! In other words we have

$$\hat{P}_a = \sum_{j,x} |a,j,x\rangle\langle a,j,x| = \frac{n_a}{N} \sum_g \chi^*_a(g) R(g) \,. \tag{9.143}$$

In summary

- P_a is a projector in the space generated by one irreducible representation, that is the space of all $|a, j, x\rangle$ for all j and x. That is, on the Hilbert space $\mathcal{H}_a = \oplus_x \mathcal{H}_{a,x}$.
- Π^a_{jj} is a projector on the subspace $|a, j, x\rangle$ for all x, but with a fixed j (that is , one of the dimension of the representation).
- Π^a_{kj} is a generalised projector.

As ever let's quickly convince ourselves that this does actually work by looking at a quick example. Consider the Pauli group $G = \{\pm(i)\sigma_x, \pm(i)\sigma_y, \pm(i)\sigma_z, \pm(i)I\}$. We first note that all Pauli's are traceless and so their contribution vanishes leaving us with only the contribution from the identity terms. Thus we have:

$$\begin{aligned} P_a &= \frac{2}{16}(\mathrm{Tr}(I)^* I + \mathrm{Tr}(-I)^*(-I) + \mathrm{Tr}(iI)^*(iI) + \mathrm{Tr}(-iI)^*(-iI)) \\ &= \frac{4}{16}(I + I + I + I) = I\,, \end{aligned} \tag{9.144}$$

as expected. (Note the similarities between this and the group averaging results in Section 9.8).

9.10.14 Ammonia Molecule Example

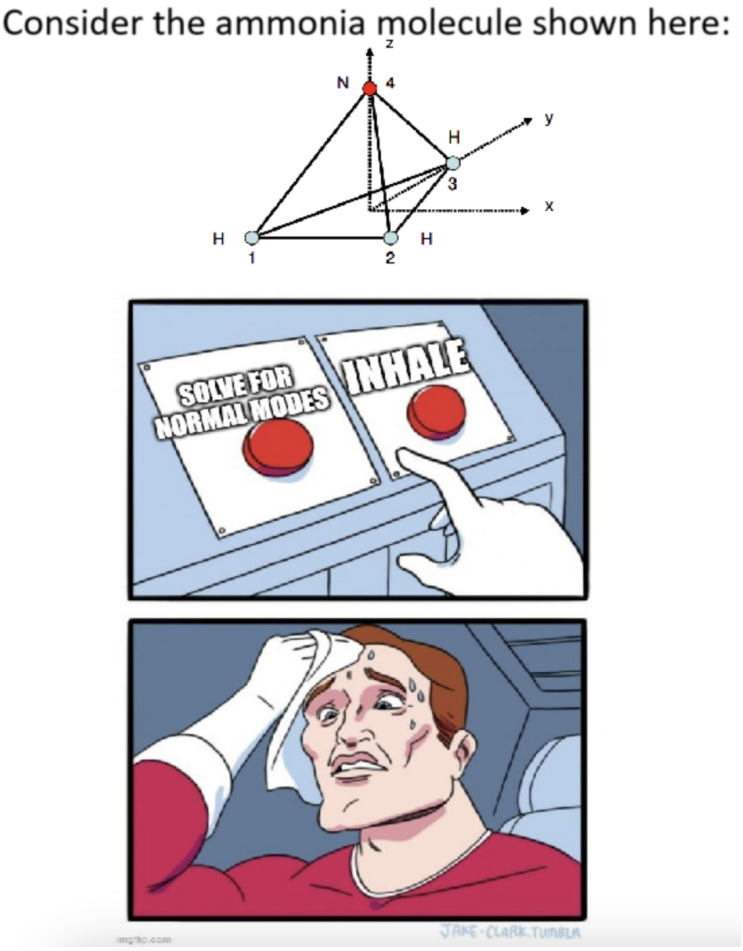

Figure 9.11: Credit: Stefan Visnjic

Let us end by tying everything together with the example of the vibrations modes of the ammonia molecule (NH_3). This will be a classical treatment; however, the lessons carry over to

quantum problems. For a more detailed introduction to the symmetry properties of the ammonia molecule see Chapter 1 of Vincenzo Savona's notes. You will also have the pleasure of working through this example in all its gory details in the problem sheet this week. In fact, if you want to take a stab at that problem sheet without any hints, stop reading now and have a go at it first. However, there's quite a bit to put together so I thought this week I'd use the notes to talk you through it.

The ammonia molecule consists of three hydrogen atoms arranged in a triangle and one nitrogen atom located on the vertical axis passing through the center of the triangle (see Figure 9.12). In molecular physics, it is known that for small displacements from the equilibrium positions, the restoring forces on the four atoms are proportional to the displacements. The molecule behaves as a system of coupled harmonic oscillators with 12 degrees of freedom (three spatial coordinates for each atom). Let's denote $\mathbf{R}_1$, $\mathbf{R}_2$, $\mathbf{R}_3$, and $\mathbf{R}_4$ as the coordinates of the three hydrogen atoms and the nitrogen atom. If the equilibrium positions of the four atoms are $\mathbf{R}_j^{(0)}$, where $j = 1, \ldots, 4$, then the displacement vectors are given by $\mathbf{u}_j = \mathbf{R}_j - \mathbf{R}_j^{(0)}$. Let m_H and m_N be the masses of the hydrogen and nitrogen atoms, respectively.

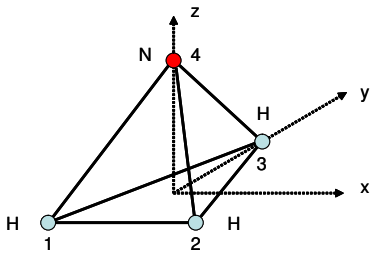

Figure 9.12: Scheme of the NH_3 molecule. In the figure, you can also see the numbering of the four atoms and the choice of the reference frame.

The symmetries of ammonia correspond to the symmetry group of a triangle in 2D. That is, we spot immediately that the relevant symmetry group here is our favourite C3v with group elements: $e, c_+, c_-, \sigma, \sigma', \sigma''$ (i.e., identity, rotations by $\pm 2\pi/3$ and reflections in each of the axis of the triangle). So what is the representation of C3v on the 12 dimensional space spanned by $\boldsymbol{u} := (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4)$ corresponding to the spacial displacements of the atoms that ammonia is made up of?

Well the representation of the identity is easy that's just:

$$R(e) = \mathbb{1}_{12\times 12} = \begin{pmatrix} \mathbb{1} & 0 & 0 & 0 \\ 0 & \mathbb{1} & 0 & 0 \\ 0 & 0 & \mathbb{1} & 0 \\ 0 & 0 & 0 & \mathbb{1} \end{pmatrix} \tag{9.145}$$

The c_+ rotation by $2\pi/3$ cyclically rotates molecules 1,2, and 3 (i.e., sends molecule 1 to 3, 2 to 1 and 3 to 2) and corresponds to a $2\pi/3$ rotation about the z axis in the x, y plane. The rotation

c_- is just the converse of this. Therefore we have:

$$R(c_+) = \begin{pmatrix} 0 & 0 & S & 0 \\ S & 0 & 0 & 0 \\ 0 & S & 0 & 0 \\ 0 & 0 & 0 & S \end{pmatrix} \quad \text{with} \quad S = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$R(c_-) = \begin{pmatrix} 0 & S^{-1} & 0 & 0 \\ 0 & 0 & S^{-1} & 0 \\ S^{-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & S^{-1} \end{pmatrix} \quad \text{with} \quad S^{-1} = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{9.146}$$

The σ reflection around the y axis is also easy to spot. This just switches the positions of molecules 1 and 2 and sends $-x$ to x (and vice versa) (see Fig. 9.12) and leaves all other coordinates invariant. Similar analysis can be applied to the σ_2 and σ_3 reflections. Thus we have:

$$R(\sigma) = \begin{pmatrix} 0 & M_1 & 0 & 0 \\ M_1 & 0 & 0 & 0 \\ 0 & 0 & M_1 & 0 \\ 0 & 0 & 0 & M_1 \end{pmatrix} \text{ with } M_1 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$R(\sigma') = \begin{pmatrix} M_2 & 0 & 0 & 0 \\ 0 & 0 & M_2 & 0 \\ 0 & M_2 & 0 & 0 \\ 0 & 0 & 0 & M_2 \end{pmatrix} \text{ with } M_2 = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{9.147}$$
$$R(\sigma'') = \begin{pmatrix} 0 & 0 & M_3 & 0 \\ 0 & M_3 & 0 & 0 \\ M_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & M_3 \end{pmatrix} \text{ with } M_3 = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Ok so now we have a representation on the 12d space spanned by the coordinates of the displacements of the ammonia molecule. We expect that this 12d representation is reducible. Indeed we know this as we've already seen that there are 3 irreps of C3v in Section 9.9.1, 2 1D irreps and 1 2D irrep. To save you flicking back I'll just copy them down to here:

The trivial 1D irrep:

$$R_1(e) = 1, R_1(c_+) = 1, R_1(c_-) = 1, R_1(\sigma) = 1, R_1(\sigma') = 1, R_1(\sigma'') = 1 \tag{9.148}$$

The 1D sign irrep:

$$R_2(e) = 1, R_2(c_+) = 1, R_2(c_-) = 1, R_2(\sigma) = -1, R_2(\sigma') = -1, R_2(\sigma'') = -1 \tag{9.149}$$

The 2D irrep:

$$R_3(e) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
$$R_3(c_+) = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \; R_3(c_-) = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \tag{9.150}$$
$$R_3(\sigma) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \; R_3(\sigma') = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \; R_3(\sigma'') = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$$

	e	$2c_+$	$3\sigma_v$
R_1	1	1	1
R_2	1	1	-1
R_3	2	-1	0
R	12	0	2

Table 9.2: Character table for point group C3v.

The corresponding C3v character table is shown in Table 9.10.14.

So how do we write our 12d rep in terms of these irreps? Well we can use Theorem 9.9.3 to compute the number of times each of these irreps appears in our rep. That is, we can use

$$b_a = \frac{1}{N}\sum_\mu n_\mu \chi_a^*(C_\mu)\chi_R(C_\mu) \tag{9.151}$$

where b_a is the degeneracy factor of representation a, n_μ is the number of group elements in conjugacy class μ and N is the order of the group. Thus we have

$$\begin{aligned} b_1 &= \frac{1}{6}(1\times 1\times 12 + 2\times 1\times 0 + 3\times 1\times 2) = 3 \\ b_2 &= \frac{1}{6}(1\times 1\times 12 + 2\times 1\times 0 - 3\times 1\times 2) = 1 \\ b_3 &= \frac{1}{6}(1\times 2\times 12 - 2\times 1\times 0 + 3\times 0\times 2) = 4 \end{aligned}$$

And thus, we can express R in terms of the irreducible representations of C_{3v} as follows:

$$R = 3R_1 \oplus R_2 \oplus 4R_3. \tag{9.152}$$

We now know how to compose R into irreps. But we do not yet know the basis to do so. That is, we need to look for a basis where 3×1 vectors are invariant under R_1, where 1×1 vectors are invariant under R_2, and 4×2 vectors are invariant under R_3 [25]. To achieve this, it suffices to choose an arbitrary basis (we choose for simplicity $\boldsymbol{v}_i = \hat{\boldsymbol{e}}_i$ where $\hat{\boldsymbol{e}}_i$ is a normalized vector with the i-th entry being the only non-zero entry) and apply various projectors.

Recall from Eq. (9.138) that a projector onto a basis state of an irrep takes the form:

$$\hat{\Pi}^b_{kj} = \frac{n_b}{N}\sum_g [R_b(g)]^*_{kj} R(g) \tag{9.153}$$

where n_b is the dimension of the representation b and N is the order of the group. That is, we have and we have

$$\hat{\Pi}^a_{kj}\,|a,j,x\rangle = |a,k,x\rangle \tag{9.154}$$

$$\hat{\Pi}^a_{kj}\,|b,j',x\rangle = 0 \quad \text{otherwise} \tag{9.155}$$

where $|a,k,x\rangle$ is a basis for the reduced representation. So if we apply Π^b_{kj} on an arbitrary state and get a non-zero vector, we are left with a (non-normalised) basis state. If we get a set of these we can create an orthonormal basis via the gram-schmidt procedure.

[25] By comparison, remember the 2 fold tensor representation of SU(2) decomposed into a direct sum of irreps of irreps on SU(3) and SU(1) *in the Bell basis*. We currently know that $R = 3R_1 \oplus R_2 \oplus 4R_3$ in some basis but we do not know which yet. In our warm up example the three Bell states $\{|\psi_+\rangle, |\phi_+\rangle, |\phi_-\rangle\}$ and

Let us start by constructing the projector corresponding to R_1:

$$\hat{\Pi}^1_{11} = \frac{1}{6}\sum_g [R_1(g)]^*_{11} R(g) = \frac{1}{6}(R(e) + R(c_+) + R(c_1) + R(\sigma) + R(\sigma') + R(\sigma'')) \,. \tag{9.156}$$

Thus we have:

$$\hat{\Pi}^1_{11} = \frac{1}{6}\begin{pmatrix} \mathbb{1} + M_2 & S^{-1} + M_1 & S + M_3 & 0 \\ S + M_1 & \mathbb{1} + M_3 & S^{-1} + M_2 & 0 \\ S^{-1} + M_3 & S + M_2 & \mathbb{1} + M_1 & 0 \\ 0 & 0 & 0 & S^{(1)} \end{pmatrix} \tag{9.157}$$

where we have defined $S^{(1)} = \mathbb{1} + M_1 + M_2 + M_3 + S + S^{-1}$. Similarly for R_2 we have

$$\hat{\Pi}^2_{11} = \frac{1}{6}\sum_g [R_2(g)]^*_{11} R(g) = \frac{1}{6}(R(e) + R(c_+) + R(c_1) - R(\sigma) - R(\sigma') - R(\sigma'')) \,. \tag{9.158}$$

so we have the same as above but get a minus sign in front of each of the M_i terms. That is,

$$\hat{\Pi}^2_{11} = \frac{1}{6}\begin{pmatrix} \mathbb{1} + M_2 & S^{-1} - M_1 & S - M_3 & 0 \\ S - M_1 & \mathbb{1} - M_3 & S^{-1} - M_2 & 0 \\ S^{-1} - M_3 & S - M_2 & \mathbb{1} - M_1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{9.159}$$

where we note that $\mathbb{1} - M_1 - M_2 - M_3 + S + S^{-1} = 0$. I'll leave it up to you to compute $\hat{\Pi}^3_{11}$ yourself.

We can now use the projectors to find a basis for each of the irreps. Let us start with R_1. We can find the first 3 basis vectors by evaluating $\boldsymbol{u} \equiv \hat{\Pi}^1_{11}\boldsymbol{v}$ and to give a set of 3 linearly independent vectors $\boldsymbol{u}$ in this basis. One possible choice (not necessarily unique, also dependent on the basis $\boldsymbol{v}_i$) is to select: $\hat{\Pi}^1_{11}\boldsymbol{v}_1$, $\hat{\Pi}^1_{11}\boldsymbol{v}_3$, and $\hat{\Pi}^1_{11}\boldsymbol{v}_{12}$. Then each of these vectors needs to be orthonormalized using, for example, a Gram-Schmidt algorithm. We'll use $\hat{\boldsymbol{u}}$ to denote the vector of the constructed basis after they have been orthonormalized. It's an iterative procedure, before adding a vector to the basis, it needs to be orthonormalized with respect to those already in the basis.

Let's see what this looks like for R_1. We start by evaluating

$$\begin{aligned} \boldsymbol{u}_{1,1} = \hat{\Pi}^1_{11}\boldsymbol{v}_1 &= \frac{1}{6}\begin{pmatrix} \mathbb{1} + M_2 & S^{-1} + M_1 & S + M_3 & 0 \\ S + M_1 & \mathbb{1} + M_3 & S^{-1} + M_2 & 0 \\ S^{-1} + M_3 & S + M_2 & \mathbb{1} + M_1 & 0 \\ 0 & 0 & 0 & S^{(1)} \end{pmatrix}\boldsymbol{v}_1 \\ &= \left(\frac{1}{4}, \frac{\sqrt{3}}{12}, 0, -\frac{1}{4}, \frac{\sqrt{3}}{12}, 0, 0, -\frac{\sqrt{3}}{6}, 0, 0, 0, 0\right) \end{aligned} \tag{9.160}$$

which if we normalise gives

$$\hat{\boldsymbol{u}}_{1,1} = \left(\frac{1}{2}, \frac{1}{2\sqrt{3}}, 0, -\frac{1}{2}, \frac{1}{2\sqrt{3}}, 0, 0, -\frac{1}{\sqrt{3}}, 0, 0, 0, 0\right) . \tag{9.161}$$

Let's do another one

$$\boldsymbol{u}_{1,3} \equiv \hat{\Pi}^1_{11}\boldsymbol{v}_{12} = (0,0,0,0,0,0,0,0,0,0,0,1) = \hat{\boldsymbol{u}}_{1,3} \tag{9.162}$$

which conveniently is already normalised and orthogonal to $\hat{\boldsymbol{u}}_{1,1}$. You can similarly generate a third one as

$$\boldsymbol{u}_{1,2} \equiv \hat{\Pi}^1_{11}\boldsymbol{v}_3 \qquad \rightarrow \qquad \hat{\boldsymbol{u}}_{1,2} = \left(0,0,\frac{1}{\sqrt{3}},0,0,\frac{1}{\sqrt{3}},0,0,\frac{1}{\sqrt{3}},0,0,0\right) \tag{9.163}$$

where in this case you need to explicitly apply Gram Schmidt to ensure $\hat{\boldsymbol{u}}_{1,2}$ is normalised and orthogonal to $\hat{\boldsymbol{u}}_{1,1}$ and $\hat{\boldsymbol{u}}_{1,3}$. I've given you the answer above but please do work through and check you get that yourself.

The basis for R_2 is simple as its only 1D. For example, we can just do $\hat{\Pi}^2_{11}\boldsymbol{v}_1$ to get

$$\boldsymbol{u}_{2,1} \equiv \hat{\Pi}^2_{11}\boldsymbol{v}_1 \qquad \rightarrow \qquad \hat{\boldsymbol{u}}_{2,1} = \left(\frac{1}{2\sqrt{3}},-\frac{1}{2},0,\frac{1}{2\sqrt{3}},\frac{1}{2},0,-\frac{1}{\sqrt{3}},0,0,0,0,0\right). \tag{9.164}$$

Finally we come to the basis for R_3. This is more subtle. For R_3, we need to find four pairs of invariant vectors that live in the same invariant subspace. We can do this by creating the first vector using the same procedure as above, i.e. as $\boldsymbol{u} \equiv \Pi^{(3)}_{11}\boldsymbol{v}$ for some $\boldsymbol{v}$. On normalizing we'll have $\hat{\boldsymbol{u}} = |3,1,x\rangle$. Then to find another vector in the same invariant subspace we can take $\boldsymbol{u}' \equiv \Pi^{(3)}_{21}\hat{\boldsymbol{u}}$ such that we have

$$\hat{\Pi}^3_{21}|3,1,x\rangle \;=\; |3,2,x\rangle\,. \tag{9.165}$$

This way we can be sure that $\hat{\boldsymbol{u}}$ and $\hat{\boldsymbol{u}}'$ leave the same invariant subspace. It's a little long to do but works. I'll leave the fun* of doing so to you.

Putting it all together, we end up with a set of ortho-normal vectors. We can then use these to construct the unitary that transforms into the basis in which R decomposes into irreps:

$$\boldsymbol{U} = (\hat{\boldsymbol{u}}_{1,1},\hat{\boldsymbol{u}}_{1,2},\hat{\boldsymbol{u}}_{1,3},\hat{\boldsymbol{u}}_{2,1},\hat{\boldsymbol{u}}_{3,1},\hat{\boldsymbol{u}}_{3,2},\hat{\boldsymbol{u}}_{3,4},\hat{\boldsymbol{u}}_{3,3},\hat{\boldsymbol{u}}'_{3,1},\hat{\boldsymbol{u}}'_{3,2},\hat{\boldsymbol{u}}'_{3,4},\hat{\boldsymbol{u}}'_{3,3}) \tag{9.166}$$

Why is this good to know? Well from Schur's lemmas we know that we can use the irrep structure to block diagonalize any operator that commutes with all representations of elements of the group. Thus we see immediately that we can block diagonalize any 12d operator that commutes with $R(g)$ for all g.

Let's look at an example of this. To realistically describe the harmonic modes of the ammonia, a precise parametrization of the elastic constants would be necessary. In general, we cannot express the harmonic force on an atom as the sum of harmonic forces exerted by the other atoms because the harmonic constant for the force between two atoms will be influenced by the presence of the other atoms. However, in the context of this exercise, we can introduce a highly simplified model without fear, which allows us to familiarize ourselves with the symmetry properties. We will assume that the system is simply characterized by two harmonic constants: k_{HH} for the restoring force between two hydrogen atoms and k_{NH} for the force between a hydrogen atom and the nitrogen atom. We have made a strong approximation by assuming that the harmonic force between two atoms is isotropic.

Once the masses and elastic constants are given, we can write the potential energy as follows:

$$\begin{aligned} V(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3,\mathbf{u}_4) \;&=\; \frac{1}{2}k_{HH}\left[(\mathbf{u}_1-\mathbf{u}_2)^2+(\mathbf{u}_1-\mathbf{u}_3)^2+(\mathbf{u}_2-\mathbf{u}_3)^2\right] \\ &+\; \frac{1}{2}k_{NH}\left[(\mathbf{u}_1-\mathbf{u}_4)^2+(\mathbf{u}_2-\mathbf{u}_4)^2+(\mathbf{u}_3-\mathbf{u}_4)^2\right]. \end{aligned} \tag{9.167}$$

The force acting on a given particle is obtained from the gradient of this potential with respect to the corresponding displacement variable:

$$\mathbf{F}_j = m_j \frac{\partial^2 \mathbf{u}_j}{\partial t^2} = -\frac{\partial V}{\partial \mathbf{u}_j}, \tag{9.168}$$

which allows us to write the equations of motion for the system:

$$\begin{aligned}
m_H \frac{\partial^2 \mathbf{u}_1}{\partial t^2} &= -k_{HH}(\mathbf{u}_1 - \mathbf{u}_2) - k_{HH}(\mathbf{u}_1 - \mathbf{u}_3) - k_{NH}(\mathbf{u}_1 - \mathbf{u}_4)\,, \\
m_H \frac{\partial^2 \mathbf{u}_2}{\partial t^2} &= -k_{HH}(\mathbf{u}_2 - \mathbf{u}_1) - k_{HH}(\mathbf{u}_2 - \mathbf{u}_3) - k_{NH}(\mathbf{u}_2 - \mathbf{u}_4)\,, \\
m_H \frac{\partial^2 \mathbf{u}_3}{\partial t^2} &= -k_{HH}(\mathbf{u}_3 - \mathbf{u}_1) - k_{HH}(\mathbf{u}_3 - \mathbf{u}_2) - k_{NH}(\mathbf{u}_3 - \mathbf{u}_4)\,, \\
m_N \frac{\partial^2 \mathbf{u}_4}{\partial t^2} &= -k_{NH}(\mathbf{u}_4 - \mathbf{u}_1) - k_{NH}(\mathbf{u}_4 - \mathbf{u}_2) - k_{NH}(\mathbf{u}_4 - \mathbf{u}_3)\,.
\end{aligned} \tag{9.169}$$

In this simplified notation, it is implied that the variables $\mathbf{u}_j(t)$ depend on time. Such a system of coupled oscillators is characterized by "normal modes." A normal mode is a specific solution to the equations (9.169) where the 12 degrees of freedom depend on time according to the same harmonic law:

$$\mathbf{u}_j(t) = \mathbf{u}_j^{(0)} \sin(\omega t)\,. \tag{9.170}$$

Here, $\mathbf{u}_j^{(0)}$ is a constant vector. By substituting the solution (9.170) into the set of equations (9.169), we obtain:

$$\begin{aligned}
\omega^2 \mathbf{u}_1^{(0)} &= \frac{1}{m_H}\left[k_{HH}(\mathbf{u}_1^{(0)} - \mathbf{u}_2^{(0)}) + k_{HH}(\mathbf{u}_1^{(0)} - \mathbf{u}_3^{(0)}) + k_{NH}(\mathbf{u}_1^{(0)} - \mathbf{u}_4^{(0)})\right], \\
\omega^2 \mathbf{u}_2^{(0)} &= \frac{1}{m_H}\left[k_{HH}(\mathbf{u}_2^{(0)} - \mathbf{u}_1^{(0)}) + k_{HH}(\mathbf{u}_2^{(0)} - \mathbf{u}_3^{(0)}) + k_{NH}(\mathbf{u}_2^{(0)} - \mathbf{u}_4^{(0)})\right], \\
\omega^2 \mathbf{u}_3^{(0)} &= \frac{1}{m_H}\left[k_{HH}(\mathbf{u}_3^{(0)} - \mathbf{u}_1^{(0)}) + k_{HH}(\mathbf{u}_3^{(0)} - \mathbf{u}_2^{(0)}) + k_{NH}(\mathbf{u}_3^{(0)} - \mathbf{u}_4^{(0)})\right], \\
\omega^2 \mathbf{u}_4^{(0)} &= \frac{1}{m_N}\left[k_{NH}(\mathbf{u}_4^{(0)} - \mathbf{u}_1^{(0)}) + k_{NH}(\mathbf{u}_4^{(0)} - \mathbf{u}_2^{(0)}) + k_{NH}(\mathbf{u}_4^{(0)} - \mathbf{u}_3^{(0)})\right].
\end{aligned} \tag{9.171}$$

Subsequently, to simplify the notation, we will represent $\mathbf{u}_j^{(0)}$ as simply $\mathbf{u}_j$. We can define the vector in the 12-dimensional space as:

$$\mathbf{u} = (\mathbf{u}_1;\, \mathbf{u}_2;\, \mathbf{u}_3;\, \mathbf{u}_4)\,. \tag{9.172}$$

The system of equations (9.171) can be expressed in the compact form:

$$A\mathbf{u} = \omega^2 \mathbf{u}\,, \tag{9.173}$$

Here, A is the dynamic matrix of the system, obtained straightforwardly from the form (9.171) of the equation of motion and takes the form:

$$A = \begin{pmatrix} A_{DH} & A_{HH} & A_{HH} & A_{HN} \\ A_{HH} & A_{DH} & A_{HH} & A_{HN} \\ A_{HH} & A_{HH} & A_{DH} & A_{HN} \\ -A_{DN} & -A_{DN} & -A_{DN} & 3A_{DN} \end{pmatrix} \tag{9.174}$$

where we have defined $A_{DH} = a_{DH}\mathbf{1}$, $A_{HH} = a_{HH}\mathbf{1}$, $A_{HN} = a_{HN}\mathbf{1}$ $A_{DN} = a_{DN}\mathbf{1}$ are 3×3 matrices multiples of the identity and $a_{DH} = (2k_{HH} + k_{HN})/m_H$, $a_{HH} = -k_{HH}/m_H$, $a_{HN} = -k_{HN}/m_H$, $a_{DN} = k_{HN}/m_N$.

Exercise: Show this!

To find the normal modes of the molecule we need to solve the eigenvalue problem in Eq. (9.173). That is, we want to diagonalize A. We can use our new found knowledge of the irreps of R to do this. That is, the matrix $\tilde{A} = U^\dagger A U$ (where U is the basis change we worked so hard to find earlier in Eq. (9.166)) is block-diagonal:

$$\tilde{A} = \begin{pmatrix} \tilde{A}_{1\times1} & & & & & \\ & \tilde{A}_{2\times2} & & & & \\ & & \tilde{A}_{1\times1} & & & \\ & & & \tilde{A}_{3\times3} & & \\ & & & & \tilde{A}_{1\times1} & \\ & & & & & \tilde{A}_{3\times3} \\ & & & & & & \tilde{A}_{1\times1} \end{pmatrix} \tag{9.175}$$

with :

$$\tilde{A}_{1\times1} = \frac{3k_{\mathrm{HH}} + k_{\mathrm{HN}}}{m_{\mathrm{H}}} \tag{9.176}$$

$$\tilde{A}_{2\times2} = \begin{pmatrix} \frac{k_{\mathrm{HN}}}{m_{\mathrm{H}}} & -\frac{\sqrt{3}k_{\mathrm{HN}}}{m_{\mathrm{H}}} \\ -\frac{\sqrt{3}k_{\mathrm{HN}}}{m_{\mathrm{N}}} & \frac{3k_{\mathrm{HN}}}{m_{\mathrm{N}}} \end{pmatrix} \tag{9.177}$$

$$\tilde{A}_{3\times3} = \begin{pmatrix} \frac{3k_{\mathrm{HH}}+5k_{\mathrm{HN}}}{5m_{\mathrm{H}}} & \frac{6k_{\mathrm{HH}}}{5m_{\mathrm{H}}} & -\frac{2\sqrt{\frac{3}{5}}k_{\mathrm{HN}}}{m_{\mathrm{H}}} \\ \frac{6k_{\mathrm{HH}}}{5m_{\mathrm{H}}} & \frac{12k_{\mathrm{HH}}+5k_{\mathrm{HN}}}{5m_{\mathrm{H}}} & \frac{\sqrt{\frac{3}{5}}k_{\mathrm{HN}}}{m_{\mathrm{H}}} \\ -\frac{2\sqrt{\frac{3}{5}}k_{\mathrm{HN}}}{m_{\mathrm{N}}} & \frac{\sqrt{\frac{3}{5}}k_{\mathrm{HN}}}{m_{\mathrm{N}}} & \frac{3k_{\mathrm{HN}}}{m_{\mathrm{N}}} \end{pmatrix} \tag{9.178}$$

The rows in the matrix $\tilde{A}$ separate the different invariant subspaces R_i. **Exercise: Show this!**

Thus we see that the use of group theory has reduced a problem consisting of diagonalizing a 12×12 matrix to a problem requiring the diagonalization of a 2×2 matrix and a 3×3 matrix, which is much simpler!

Now we can use some physical arguments to intuitively understand what will happen if we fully solve the problem. For example, we noticed that the system is invariant under translation, so a translation along z should not cost any energy. To see this, let's look in the subspace related to Γ_1 and search for a null mode. Clearly, $\tilde{A}_{1\times1}$ is not zero, so $\tilde{A}_{2\times2}$ must have a

zero eigenvalue. Moreover, by cleverly combining $\hat{\boldsymbol{u}}_{1,2}$ and $\hat{\boldsymbol{u}}_{1,3}$, we can generate the vector $(0, 0, 1/2, 0, 0, 1/2, 0, 0, 1/2, 0, 0, 1/2)$, which is effectively a translation of each atom along z. To find the other two null modes related to translations along x and y, we need to look in Γ_3. We will find that $\tilde{A}_{3\times3}$ must also have a zero eigenvalue.

Figure 9.13: Credit: L'heure est grave