
Notes on Physical Chemistry

Notes I took based loosely on the Natural Sciences curriculum, and then some

Compiled By

ZIJUN ZHAO

*Christ's College
University of Cambridge*

Summer 2019 - Present

Contents

| | | |
|----------|---|----------|
| I | Quantum mechanics and applications | 1 |
| 1 | Quantum Mechanics | 3 |
| 1.1 | Introduction | 3 |
| 1.1.1 | Normalisation | 3 |
| 1.1.2 | Momentum | 5 |
| 1.1.3 | The uncertainty principle | 7 |
| 1.2 | Time-independent Schrödinger equation | 10 |
| 1.2.1 | Stationary states | 10 |
| 1.2.2 | The infinite square well | 14 |
| 1.2.3 | The harmonic oscillator | 20 |
| 1.2.3.1 | Ladder operator | 21 |
| 1.2.3.2 | Analytic Method | 31 |
| 1.2.3.3 | The Morse potential | 38 |
| 1.2.4 | The free particle | 40 |
| 1.2.5 | The delta-function potential | 46 |
| 1.2.5.1 | Bound states and scattering states | 46 |
| 1.2.5.2 | The Dirac delta function | 46 |

| | | |
|---------|--|----|
| 1.2.5.3 | Delta potential: bound states | 48 |
| 1.2.5.4 | Delta potential: scattering states | 50 |
| 1.3 | Formalism | 52 |
| 1.3.1 | Hilbert space | 52 |
| 1.3.2 | Observables | 53 |
| 1.3.2.1 | Hermitian operators | 53 |
| 1.3.2.2 | Determinate states | 55 |
| 1.3.3 | Eigenfunctions of a hermitian operator | 56 |
| 1.3.3.1 | Discrete spectra | 56 |
| 1.3.3.2 | Continuous spectra | 57 |
| 1.3.4 | Generalised statistical interpretation | 59 |
| 1.3.5 | The uncertainty principle | 61 |
| 1.3.5.1 | Proof of the generalised uncertainty principle | 61 |
| 1.3.5.2 | The minimum-uncertainty wave packet | 65 |
| 1.3.5.3 | The energy-time uncertainty principle | 66 |
| 1.3.6 | Dirac Notation | 70 |
| 1.4 | Quantum mechanics in three dimensions | 72 |
| 1.4.1 | Schrödinger equation in 3D with non-Coulombic potentials | 72 |
| 1.4.1.1 | Preparatory work | 72 |
| 1.4.1.2 | The angular equation | 74 |
| 1.4.1.3 | The radial equation | 80 |
| 1.4.2 | The hydrogen atom | 82 |
| 1.4.2.1 | Setting up the radial equation | 82 |
| 1.4.2.2 | Solving the radial equation | 83 |

| | | |
|----------|---|------------|
| 1.4.2.3 | Larger nuclei | 87 |
| 1.4.2.4 | Discussion of hydrogen atomic orbitals | 88 |
| 1.4.3 | Angular momentum | 91 |
| 1.4.3.1 | Operator properties | 91 |
| 1.4.3.2 | Angular momentum ladder operators | 92 |
| 1.4.3.3 | Eigenfunctions | 97 |
| 1.4.4 | The rigid rotor | 100 |
| 1.4.5 | Spin | 101 |
| 1.4.5.1 | Formalism | 101 |
| 1.4.5.2 | Spin one half | 102 |
| 1.4.5.3 | Addition of angular momenta - the intuitive way | 105 |
| 1.4.5.4 | Addition of angular momenta - the ground-up way | 106 |
| 1.4.5.5 | Permitted values of total momentum | 108 |
| 1.4.5.6 | Relation between coupling schemes | 109 |
| 1.4.5.7 | Atomic term symbols | 111 |
| 1.4.5.8 | Molecular term symbols | 115 |
| 1.5 | Identical particles | 118 |
| 1.5.1 | Two particle systems | 118 |
| 1.5.1.1 | The Schrödinger equation | 118 |
| 1.5.1.2 | Centre-of-mass coordinates | 118 |
| 1.5.1.3 | Bosons and fermions | 120 |
| 1.5.1.4 | Exchange forces | 122 |
| 2 | Spectroscopy | 127 |

| | | |
|----------|---|------------|
| 2.1 | Atomic spectroscopy | 127 |
| 2.1.1 | Selection rules | 127 |
| 2.1.2 | Fine structure | 129 |
| 2.1.2.1 | Orbital and spin magnetic moments | 129 |
| 2.1.2.2 | Spin-orbit coupling | 131 |
| 2.2 | Molecular spectroscopy | 135 |
| 2.2.1 | Spectroscopic methods | 135 |
| 2.2.1.1 | IR spectroscopy | 135 |
| 2.2.1.2 | Raman spectroscopy | 136 |
| 2.2.2 | Transition frequencies | 138 |
| 2.2.2.1 | The vibrating rotor | 138 |
| 2.2.2.2 | Determination of rotational constants | 140 |
| 2.2.2.3 | Q-branch | 141 |
| 2.2.3 | Polyatomic molecules | 143 |
| 2.2.3.1 | Normal coordinates | 143 |
| 2.2.3.2 | Symmetry of vibrational wavefunctions | 144 |
| 2.2.3.3 | Symmetry selection rules for infrared | 145 |
| 2.2.4 | Electronic spectroscopy | 146 |
| 2.2.4.1 | The Franck-Condon principle | 146 |
| 3 | Approximation methods | 149 |
| 3.1 | The variational method | 150 |
| 3.1.1 | Theory | 150 |
| 3.1.2 | Examples | 151 |

| | | |
|----------|---|------------|
| 3.1.3 | Secular determinants | 155 |
| 3.1.4 | The Hückel approximations | 157 |
| 3.1.4.1 | Example 1 | 158 |
| 3.2 | Perturbation theory | 162 |
| 3.2.1 | Nondegenerate theory | 162 |
| 3.2.2 | Examples | 165 |
| 3.2.3 | Degenerate theory | 166 |
| 3.3 | The Hartree-Fock method | 172 |
| 3.3.1 | Self-consistent field method | 172 |
| 3.3.2 | Accounting for spin | 173 |
| 3.3.3 | The Hartree-Fock method | 175 |
| 4 | Group theory | 179 |
| 4.1 | The bare minimum | 179 |
| 4.1.1 | Symmetry elements and symmetry operations | 179 |
| 4.1.2 | Point groups | 180 |
| 4.1.3 | Representations | 180 |
| 4.2 | Mathematical formulation | 182 |
| 4.2.1 | Notations | 182 |
| 4.2.2 | Basic theorems | 182 |
| 4.2.3 | Irreps and orthogonality theorems | 185 |
| 4.2.4 | Reducing representation | 187 |
| 4.2.5 | Symmetry-adapted bases | 188 |
| 4.2.6 | Decomposition of direct product bases | 190 |

| | | |
|----------|--|------------|
| 4.2.7 | The full rotation group | 192 |
| 4.2.8 | Group theory and angular momenta | 195 |
| 4.3 | Applications | 197 |
| 4.3.1 | Vanishing integrals | 197 |
| 4.3.2 | Degeneracy | 198 |
| 4.3.3 | Chemical bonding | 200 |
| 4.3.4 | LCAOs | 201 |
| 4.3.5 | Molecular orbital correlation diagrams | 201 |
| 4.3.6 | Molecular vibrations | 203 |
| 4.3.7 | Particular vibrations | 205 |
| 4.3.8 | Projection operator method | 205 |
| 5 | Solid state chemistry | 209 |
| 5.1 | Preliminary | 209 |
| 5.1.1 | Bravais lattices | 209 |
| 5.1.1.1 | Definitions | 209 |
| 5.1.1.2 | Types of Bravais lattices | 210 |
| 5.1.1.3 | The cubic lattices | 211 |
| 5.1.1.4 | Lattices with a motif | 212 |
| 5.1.2 | The ionic model | 212 |
| 5.1.2.1 | The model | 212 |
| 5.1.2.2 | Radius ratios | 214 |
| 5.2 | Free-electron gas model | 216 |
| 5.2.1 | Unconstrained FEG | 216 |

| | | |
|-----------|---|------------|
| 5.2.1.1 | One-dimensional FEG | 216 |
| 5.2.1.2 | Three-dimensional FEG | 216 |
| 5.2.2 | Quantization | 217 |
| 5.2.3 | Fermi-Dirac distribution | 219 |
| 5.2.4 | Properties of the FEG | 220 |
| 5.2.4.1 | Density of states | 220 |
| 5.2.4.2 | Fermi quantities | 221 |
| 5.2.4.3 | Average energy | 222 |
| 5.2.4.4 | Heat capacity | 222 |
| 5.2.4.5 | Bulk modulus | 223 |
| 5.2.4.6 | Electrical conductivity | 224 |
| 5.2.5 | The nearly-free electron gas | 225 |
| 5.2.5.1 | Failure of the FEG and Brillouin zones | 225 |
| 5.2.5.2 | Reciprocal lattice | 226 |
| 5.2.5.3 | Boundary cases | 228 |
| 5.3 | Tight-binding model | 230 |
| 5.3.1 | Bands in one dimension | 230 |
| 5.3.1.1 | LCAO method | 230 |
| II | Thermodynamics and statistical mechanics | 233 |
| 6 | Thermodynamics | 235 |
| 6.1 | Entropy and the Boltzmann law | 235 |
| 6.2 | The fundamental thermodynamic equations | 239 |

| | | |
|---------|--|-----|
| 6.2.1 | The fundamental equations | 239 |
| 6.2.2 | The ideal gas law | 240 |
| 6.2.3 | Second law and equilibria | 242 |
| 6.2.3.1 | The second law | 242 |
| 6.3 | The first and second laws in depth | 247 |
| 6.3.1 | The First Law | 247 |
| 6.3.1.1 | Energy | 247 |
| 6.3.1.2 | Isothermal and Adiabatic Processes | 249 |
| 6.3.2 | The Second Law | 250 |
| 6.3.2.1 | Heat Engines | 250 |
| 6.3.2.2 | Clausius' theorem | 254 |
| 6.3.3 | Entropy | 256 |
| 6.4 | Mathematical methods | 258 |
| 6.4.1 | Legendre transforms | 258 |
| 6.4.2 | Euler-Maclaurin formula | 261 |
| 6.4.3 | Homogeneous functions | 262 |
| 6.5 | Free energies | 263 |
| 6.5.1 | Helmholtz free energy | 263 |
| 6.5.1.1 | Inspiration | 263 |
| 6.5.1.2 | A microscopic look at Helmholtz I: Dimerisation | 265 |
| 6.5.1.3 | A microscopic look at Helmholtz II: Polymer collapse | 266 |
| 6.5.2 | Enthalpy | 267 |
| 6.5.3 | Gibbs free energy | 268 |
| 6.5.4 | More on thermodynamic functions | 269 |

| | | |
|----------|--|------------|
| 6.5.4.1 | Functional dependencies | 269 |
| 6.5.4.2 | Non-standard states | 270 |
| 6.5.4.3 | Gibbs-Helmholtz equation | 271 |
| 6.5.5 | The Euler and Gibbs-Duhem equations | 272 |
| 6.5.5.1 | The Euler equation | 272 |
| 6.5.5.2 | The Gibbs-Duhem equation | 273 |
| 6.6 | Maxwell's relations and mixtures | 275 |
| 6.6.1 | Designing a fundamental equation | 275 |
| 6.6.2 | Maxwell's relations | 276 |
| 6.6.3 | Susceptibilities | 277 |
| 6.6.4 | Entropic and enthalpic components | 278 |
| 6.6.5 | Partial molar properties | 280 |
| 7 | Statistical mechanics | 283 |
| 7.1 | The Boltzmann distribution law | 283 |
| 7.1.1 | The canonical ensemble | 283 |
| 7.1.1.1 | Derivation | 283 |
| 7.1.2 | Derivation of Boltzmann distribution | 284 |
| 7.1.2.1 | Via a Lagrange multiplier | 284 |
| 7.1.2.2 | From the canonical distribution | 287 |
| 7.1.3 | Interpretation of partition function | 288 |
| 7.1.3.1 | Physical meaning of the partition function | 288 |
| 7.1.3.2 | Distinguishability | 288 |
| 7.1.3.3 | Gibb's paradox | 290 |

| | | |
|---------|---|-----|
| 7.1.4 | Applications of the Boltzmann distribution | 291 |
| 7.1.4.1 | Population of states | 291 |
| 7.1.4.2 | Density of states | 292 |
| 7.1.4.3 | Non-equilibrium states | 293 |
| 7.1.4.4 | Maxwell-Boltzmann distribution of velocities | 293 |
| 7.1.5 | Thermodynamic properties from partition functions | 295 |
| 7.1.5.1 | Distinguishable particles | 295 |
| 7.1.5.2 | Indistinguishable particles | 296 |
| 7.1.6 | Examples of the two-state model | 297 |
| 7.1.6.1 | The Schottky model | 297 |
| 7.1.6.2 | Curie's law of paramagnetism | 298 |
| 7.1.6.3 | Modelling a protein loop | 299 |
| 7.2 | The statistical mechanics of simple gases | 300 |
| 7.2.1 | Molecular partition function | 300 |
| 7.2.1.1 | Translational partition function | 300 |
| 7.2.1.2 | Vibrational partition function | 301 |
| 7.2.1.3 | Rotational partition function | 302 |
| 7.2.1.4 | Electronic partition function | 303 |
| 7.2.2 | Internal energy and heat capacity | 304 |
| 7.2.3 | Entropy | 306 |
| 7.2.4 | Nuclear spin partition function | 308 |
| 7.2.4.1 | Nuclear exchange | 309 |
| 7.2.4.2 | Consequences on spectroscopy | 309 |
| 7.2.4.3 | Consequences on rotational partition function | 313 |

| | | |
|---------|---|-----|
| 7.3 | Chemical equilibrium | 314 |
| 7.3.1 | Preliminaries | 314 |
| 7.3.1.1 | Chemical potential | 314 |
| 7.3.1.2 | Choice of energy zero | 315 |
| 7.3.2 | Equilibrium constants | 315 |
| 7.3.2.1 | Calculation of K_c | 315 |
| 7.3.2.2 | Calculation of K_p | 317 |
| 7.3.3 | Examples | 318 |
| 7.3.3.1 | Dissociation of a diatomic | 318 |
| 7.3.3.2 | Ionisation of atoms | 319 |
| 7.3.3.3 | Adsorption of a gas | 320 |
| 7.4 | Rate constants | 325 |
| 7.4.1 | Transition state theory | 325 |
| 7.4.1.1 | The theory | 325 |
| 7.4.1.2 | Calculation of rate constant | 326 |
| 7.4.2 | Applications of transition state theory | 327 |
| 7.4.2.1 | Rate constant calculation | 327 |
| 7.4.2.2 | Steric factor | 328 |
| 7.4.2.3 | Kinetic isotope effects | 330 |
| 7.4.3 | Thermodynamic formulation of transition state theory | 331 |
| 7.4.3.1 | Transition Gibbs energy | 331 |
| 7.4.3.2 | Relationship to the Arrhenius equation | 332 |
| 7.4.3.3 | Experimental determination of $\Delta_r S^{\circ,\ddagger}$ and $\Delta_r H^{\circ,\ddagger}$ | 335 |
| 7.4.3.4 | Volume of activation | 335 |

| | | |
|---------|---|-----|
| 7.4.3.5 | Interpretation of experimental data | 336 |
|---------|---|-----|

| | | |
|----------|------------------------|------------|
| A | Quick reference | 341 |
|----------|------------------------|------------|

Part I

Quantum mechanics and applications

Chapter 1

Quantum Mechanics

1.1 INTRODUCTION

The 1-D Schrödinger equation is given by

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V\Psi \quad (1.1.1)$$

1.1.1 Normalisation

Theorem 1.1.1.1 (Preservation of normalisation). Solutions to the Schrödinger equation automatically preserves normalisation.

Proof. We need to show that

$$\frac{d}{dt} \int_{-\infty}^{+\infty} |\Psi(x, t)|^2 dx = 0. \quad (1.1.2)$$

We do so by directly evaluating using integration by parts.

$$\frac{d}{dt} \int_{-\infty}^{+\infty} |\Psi(x, t)|^2 dx \quad (1.1.3a)$$

$$= \int_{-\infty}^{+\infty} \frac{\partial}{\partial t} (\Psi^* \Psi) dx \quad (1.1.3b)$$

$$= \int_{-\infty}^{+\infty} \left(\Psi^* \frac{\partial \Psi}{\partial t} + \Psi \frac{\partial \Psi^*}{\partial t} \right) dx, \quad (1.1.3c)$$

$$\because \frac{\partial \Psi}{\partial t} = \frac{i\hbar}{2m} \frac{\partial^2 \Psi}{\partial x^2} - \frac{i}{\hbar} V \Psi \quad (1.1.3d)$$

$$\because \frac{\partial \Psi^*}{\partial t} = \frac{-i\hbar}{2m} \frac{\partial^2 \Psi^*}{\partial x^2} + \frac{i}{\hbar} V \Psi^* \quad (1.1.3e)$$

$$\Rightarrow \text{integrand} = \Psi^* \left(\frac{i\hbar}{2m} \frac{\partial^2 \Psi}{\partial x^2} - \frac{i}{\hbar} V \Psi \right) + \Psi \left(\frac{-i\hbar}{2m} \frac{\partial^2 \Psi^*}{\partial x^2} + \frac{i}{\hbar} V \Psi^* \right) \quad (1.1.3f)$$

$$= \frac{i\hbar}{2m} \left(\Psi^* \frac{\partial^2 \Psi}{\partial x^2} - \Psi \frac{\partial^2 \Psi^*}{\partial x^2} \right) \quad (1.1.3g)$$

$$= \frac{\partial}{\partial x} \left[\frac{i\hbar}{2m} \left(\Psi^* \frac{\partial \Psi}{\partial x} - \Psi \frac{\partial \Psi^*}{\partial x} \right) \right] \quad (1.1.3h)$$

$$\Rightarrow \frac{d}{dt} \int_{-\infty}^{+\infty} |\Psi(x, t)|^2 dx \quad (1.1.3i)$$

$$= \frac{i\hbar}{2m} \left(\Psi^* \frac{\partial \Psi}{\partial x} - \Psi \frac{\partial \Psi^*}{\partial x} \right) \Big|_{-\infty}^{+\infty}, \quad (1.1.3j)$$

where Equation 1.1.3j results from the fundamental theorem of calculus. For all normalisable wavefunctions, $\Psi \rightarrow 0$ as $|x| \rightarrow \infty$, this implies

$$\frac{d}{dt} \int_{-\infty}^{+\infty} |\Psi(x, t)|^2 dx = 0. \quad (1.1.4)$$

□

Worked example 1.1.1.1. Consider the wave function

$$\Psi(x, t) = A e^{-\lambda|x|} e^{-i\omega t}, \quad (1.1.5)$$

where A , λ and ω are positive real constants.

(a) Normalise Ψ .

The time dependence cancels out in

$$\int_{-\infty}^{+\infty} \Psi^* \Psi dx, \quad (1.1.6)$$

leaving us with

$$\begin{aligned}
 & 2A^2 \int_0^\infty e^{-2\lambda x} dx \\
 &= \frac{A^2}{\lambda} \\
 &= 1 \\
 &\Rightarrow A = \sqrt{\lambda}
 \end{aligned} \tag{1.1.7}$$

We therefore have the normalised wavefunction as

$$\Psi(x, t) = \sqrt{\lambda} e^{-\lambda|x|} e^{-i\omega t}. \tag{1.1.8}$$

(b) Determine the standard deviation of x and calculate the probability that x falls in the range of $\langle x \rangle \pm \sigma$.

As the wavefunction is an even function, $\langle x \rangle = 0$.

We still need to calculate $\langle x^2 \rangle$:

$$\begin{aligned}
 & 2\lambda \int_0^\infty x^2 e^{-2\lambda x} dx \\
 &= \frac{1}{2\lambda^2} \\
 &\because \sigma^2 = \langle x^2 \rangle - \langle x \rangle^2 \\
 &\Rightarrow \sigma = \frac{1}{\sqrt{2}\lambda}
 \end{aligned} \tag{1.1.9}$$

The required probability is then calculated by

$$\begin{aligned}
 & \lambda \int_{-1/\sqrt{2}\lambda}^{1/\sqrt{2}\lambda} e^{-2\lambda x} dx \\
 &= 1 - e^{-\sqrt{2}} \\
 &= 0.7569.
 \end{aligned} \tag{1.1.10}$$

1.1.2 Momentum

Postulate 1.1.2.1. All classical dynamical variables can be expressed in terms of position and momentum.

We already know the position operator, $\hat{X} = x$, and now we will find the momentum operator. We know that

$$\langle x \rangle = \int_{-\infty}^{+\infty} x |\Psi(x, t)|^2 dx. \tag{1.1.11}$$

We postulate that

$$\langle v \rangle = \frac{d\langle x \rangle}{dt} = \int_{-\infty}^{+\infty} x \frac{\partial}{\partial x} |\Psi(x, t)|^2 dx = \frac{i\hbar}{2m} \int_{-\infty}^{+\infty} x \frac{\partial}{\partial x} \left[\frac{i\hbar}{2m} \left(\Psi^* \frac{\partial \Psi}{\partial x} - \Psi \frac{\partial \Psi^*}{\partial x} \right) \right] dx \quad (1.1.12a)$$

$$= -\frac{i\hbar}{2m} \int_{-\infty}^{+\infty} \left(\Psi^* \frac{\partial \Psi}{\partial x} - \Psi \frac{\partial \Psi^*}{\partial x} \right) dx \quad (1.1.12b)$$

$$= -\frac{i\hbar}{m} \int_{-\infty}^{+\infty} \Psi^* \frac{\partial \Psi}{\partial x} dx, \quad (1.1.12c)$$

where Equations 1.1.12b and 1.1.12c are the results of integration by parts.

Definition 1.1.2.1 (Momentum operator). We can now extract the definition of the momentum operator, \mathbf{P} , from Equation 1.1.12c:

$$\mathbf{P} \equiv \frac{\hbar}{i} \frac{\partial}{\partial x}. \quad (1.1.13)$$

Theorem 1.1.2.1 (Ehrenfest theorem). The Ehrenfest theorem states that

$$m \frac{d}{dt} \langle x \rangle = \langle p \rangle, \text{ and} \quad (1.1.14a)$$

$$\frac{d}{dt} \langle p \rangle = -\langle V'(x) \rangle. \quad (1.1.14b)$$

Equation 1.1.14a is already proven, and we give a proof to Equation 1.1.14b below.

Proof.

$$\frac{d}{dt} \langle p \rangle = \frac{d}{dt} \left(-i\hbar \int_{-\infty}^{+\infty} \Psi^* \frac{\partial \Psi}{\partial x} dx \right) \quad (1.1.15a)$$

$$= -i\hbar \int_{-\infty}^{+\infty} \frac{\partial}{\partial t} \left(\Psi^* \frac{\partial \Psi}{\partial x} \right) dx \quad (1.1.15b)$$

$$= -i\hbar \int_{-\infty}^{+\infty} \Psi_t^* \Psi_x + \Psi^* \Psi_{xt} dx \quad (1.1.15c)$$

$$= -i\hbar \int_{-\infty}^{+\infty} \frac{-i\hbar}{2m} \Psi_{xx}^* \Psi_x + \frac{i}{\hbar} V \Psi^* \Psi_x + \Psi^* \frac{\partial}{\partial x} \left(\frac{i\hbar}{2m} \Psi_{xx} - \frac{i}{\hbar} V \Psi \right) dx \quad (1.1.15d)$$

$$= -i\hbar \int_{-\infty}^{+\infty} \frac{-i\hbar}{2m} \Psi_{xx}^* \Psi_x + \frac{i}{\hbar} V \Psi^* \Psi_x + \Psi^* \left(\frac{i\hbar}{2m} \Psi_{xxx} - \frac{i}{\hbar} V_x \Psi - \frac{i}{\hbar} V \Psi_x \right) dx \quad (1.1.15e)$$

$$= -i\hbar \int_{-\infty}^{+\infty} \frac{-i\hbar}{2m} \Psi_{xx}^* \Psi_x + \Psi^* \frac{i\hbar}{2m} \Psi_{xxx} - \frac{i}{\hbar} V_x \Psi \Psi^* dx \quad (1.1.15f)$$

We attempt to reduce the expression by evaluating $\int_{-\infty}^{+\infty} \Psi_{xx}^* \Psi_x dx$ by parts twice:

$$\int_{-\infty}^{+\infty} \Psi_{xx}^* \Psi_x dx = \Psi_x^* \Psi_x \Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} \Psi_x^* \Psi_{xx} dx \quad (1.1.16a)$$

$$= 0 - \Psi_x^* \Psi_{xx} \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} \Psi_x^* \Psi_{xxx} dx \quad (1.1.16b)$$

$$= \int_{-\infty}^{+\infty} \Psi_x^* \Psi_{xxx} dx. \quad (1.1.16c)$$

This leads to the cancellation of the first two terms in the integrand in Equation 1.1.15f and we are left with the desired result:

$$\frac{d}{dt} \langle p \rangle = - \int_{-\infty}^{+\infty} \Psi^* V_x \Psi dx = - \langle V'(x) \rangle. \quad (1.1.17)$$

□

1.1.3 The uncertainty principle

Theorem 1.1.3.1 (The uncertainty principle). The uncertainty principle relates the error in position and momentum by

$$\sigma_x \sigma_p \geq \frac{\hbar}{2}, \quad (1.1.18)$$

a result we will prove only later.

Worked example 1.1.3.1.

A particle of mass m is in the state

$$\Psi(x, t) = A e^{-a[(mx^2/\hbar) + it]}, \quad (1.1.19)$$

where A and a are positive real constants.

(a) Find A .

$$A^2 \int_{-\infty}^{+\infty} \psi \psi^* dx = 1 \quad (1.1.20a)$$

$$A^2 \int_{-\infty}^{+\infty} e^{-2a(mx^2/\hbar)} dx = 1 \quad (1.1.20b)$$

$$A^2 \sqrt{\frac{\pi \hbar}{2ma}} = 1 \quad (1.1.20c)$$

$$A = \left(\frac{2ma}{\pi \hbar} \right)^{1/4}. \quad (1.1.20d)$$

(b) For what potential energy function $V(x)$ does Ψ satisfy the Schrödinger equation ?

We have the spatial wave function

$$\psi(x) = \left(\frac{2ma}{\pi \hbar} \right)^{1/4} e^{-max^2/\hbar}, \quad (1.1.21)$$

which when differentiated twice gives

$$\left(\frac{2ma}{\pi \hbar} \right)^{1/4} \left(\frac{4m^2 a^2}{\hbar^2} x^2 - \frac{2ma}{\hbar} \right) e^{-max^2/\hbar} = \left(\frac{4m^2 a^2}{\hbar^2} x^2 - \frac{2ma}{\hbar} \right) \psi, \quad (1.1.22)$$

and the Schrödinger equation gives

$$\left[\left(\hbar a - 2ma^2 x^2 \right) + V \right] \psi = E \psi \quad (1.1.23)$$

The energy is found by

$$i\hbar \frac{d\phi}{dt} = E\phi \quad (1.1.24a)$$

$$E = \hbar a, \quad (1.1.24b)$$

with $\phi = e^{-iat}$. The potential must then be

$$V(x) = 2ma^2 x^2, \quad (1.1.25)$$

the harmonic oscillator potential.

(c) Calculate the expectation values of x , x^2 , p , and p^2 .

$x\phi$ and $\frac{\partial \phi}{\partial x}$ are odd functions and as such

$$\langle x \rangle = \langle p \rangle = 0. \quad (1.1.26)$$

Letting $k = ma/\hbar$

$$\langle x^2 \rangle = \int_{-\infty}^{+\infty} x^2 \psi^2 dx \quad (1.1.27a)$$

$$= \left(\frac{2ma}{\pi\hbar} \right)^{1/2} \int_{-\infty}^{+\infty} x^2 e^{-kx^2} dx \quad (1.1.27b)$$

$$= -\frac{x}{2k} e^{-kx^2} \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} \frac{1}{2k} e^{-kx^2} dx \quad (1.1.27c)$$

$$= 0 + \frac{1}{2k} \sqrt{\frac{\pi}{k}} \cdot \left(\frac{2ma}{\pi\hbar} \right)^{1/2} \quad (1.1.27d)$$

$$= \frac{\hbar}{4ma}. \quad (1.1.27e)$$

$$\langle p^2 \rangle = -\hbar^2 \int_{-\infty}^{+\infty} \psi \frac{d^2\psi}{dx^2} dx \quad (1.1.28a)$$

$$= -\hbar^2 \int_{-\infty}^{+\infty} \left(\frac{4m^2 a^2}{\hbar^2} x^2 - \frac{2ma}{\hbar} \right) \psi^2 dx \quad (1.1.28b)$$

$$= -4m^2 a^2 \langle x^2 \rangle + 2\hbar ma \int_{-\infty}^{+\infty} \psi^2 dx \quad (1.1.28c)$$

$$= \hbar ma. \quad (1.1.28d)$$

Standard deviations are given by

$$\sigma_x = \sqrt{\langle x^2 \rangle - \langle x \rangle^2} = \frac{1}{2} \sqrt{\frac{\hbar}{ma}} \quad (1.1.29a)$$

$$\sigma_p = \sqrt{\langle p^2 \rangle - \langle p \rangle^2} = \sqrt{\hbar ma}, \quad (1.1.29b)$$

with the product of uncertainties

$$\sigma_x \sigma_p = \frac{\hbar}{2}. \quad (1.1.30)$$

In this case, the uncertainty is as small as possible.

1.2 TIME-INDEPENDENT SCHRÖDINGER EQUATION

1.2.1 Stationary states

We solve the Schrödinger equation of the form

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V\Psi \quad (1.2.1)$$

For a general potential $V(x)$, independent of t . We solve the partial differential equation by separation of variables:

$$\Psi(x, t) = \psi(x)\phi(t). \quad (1.2.2)$$

Substituting into the original equation we have

$$i\hbar \frac{1}{\phi} \frac{d\phi}{dt} = -\frac{\hbar^2}{2m} \frac{1}{\psi} \frac{d^2\psi}{dx^2} + V. \quad (1.2.3)$$

So we have

$$i\hbar \frac{1}{\phi} \frac{d\phi}{dt} = E \quad (1.2.4a)$$

$$-\frac{\hbar^2}{2m} \frac{1}{\psi} \frac{d^2\psi}{dx^2} + V\psi = E, \quad (1.2.4b)$$

or,

$$\frac{d\phi}{dt} = -\frac{iE}{\hbar} \phi \quad (1.2.5a)$$

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + V\psi = E\psi. \quad (1.2.5b)$$

Solving Equation 1.2.5a gives

$$\phi(t) = e^{-iEt/\hbar}. \quad (1.2.6)$$

Equation 1.2.5b is the *time-independent Schrödinger equation* and solving it requires that $V(x)$ be specified.

Property 1.2.1.1 (Stationary states). Stationary states have the following properties:

- *All expectation value is constant.* This means that, from Equation 1.1.14a,

$$\langle p \rangle = 0. \quad (1.2.7)$$

- *Total energy is definite.* Proof follows.

Proof.

$$H(x, p) = \frac{p^2}{2m} + V(x). \quad (1.2.8)$$

The corresponding quantum mechanical Hamiltonian operator is given by canonical substitution $p \rightarrow (\hbar/i)(\partial/\partial x)$:

$$H = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x). \quad (1.2.9)$$

We can then write the Schrödinger equation as

$$H\psi = E\psi, \quad (1.2.10)$$

So the expectation value of the total energy is

$$\langle H \rangle = \int_{-\infty}^{+\infty} \psi^* H\psi \, dx = E \int_{-\infty}^{+\infty} \psi^2 \, dx = E, \quad (1.2.11)$$

and

$$\langle H^2 \rangle = \int_{-\infty}^{+\infty} \psi^* H^2\psi \, dx = E^2 \int_{-\infty}^{+\infty} \psi^2 \, dx = E^2. \quad (1.2.12)$$

The variance of H is calculated by

$$\sigma_H^2 = \langle H^2 \rangle - \langle H \rangle^2 = 0 \quad (1.2.13)$$

Hence, every measurement of the total energy is certain to return the value E . \square

Corollary 1.2.1.1. As a result of the second property above, we can simplify the calculation of $\langle H \rangle$ by using the *intial wavefunction* $\Psi(x, 0)$ which usually have a simpler form than the complete wavefunction.

Theorem 1.2.1.1 (Particle-in-a-box general solution).

The general solution is a linear combination of separable solutions:

$$\Psi(x, t) = \sum_{n=1}^{\infty} c_n \psi_n e^{-iE_n t/\hbar}, \quad (1.2.14)$$

which is a *Fourier series*, whose coefficients can be determined by initial condition.

It is worth stressing that Property 1.2.1.1 *does not* apply to the general solution, because the exponential time dependences do not cancel out due to different energies, but rather individual summation terms, the stationary states.

Worked example 1.2.1.1. Suppose a particle starts out in a linear combination of just two *real* stationary states, *i.e.*, all other c_n 's are 0:

$$\Psi(x, 0) = c_1 \psi_1 + c_2 \psi_2. \quad (1.2.15)$$

What is the complete wavefunction $\Psi(x, t)$ at subsequent times? Find the probability density function and describe its motion.

Solution: The first part is trivial: just stick the time dependence onto the spatial wavefunction:

$$\Psi(x, t) = c_1 \psi_1 e^{-iE_1 t/\hbar} + c_2 \psi_2 e^{-iE_2 t/\hbar}. \quad (1.2.16)$$

So

$$|\Psi(x, t)|^2 = \Psi \Psi^* \quad (1.2.17a)$$

$$= c_1^2 \psi_1^2 + c_2^2 \psi_2^2 + 2c_1 c_2 \psi_1 \psi_2 \cos[(E_2 - E_1)t/\hbar]. \quad (1.2.17b)$$

It oscillates sinusoidally at $\omega = (E_2 - E_1)/\hbar$, so clearly the complete solution is *not* a stationary state.

We now introduce three seemingly trivial but oft-invoked theorems.

Theorem 1.2.1.2 (Energy must be real).

For all normalisable solutions, the separation constant E must be real.

Proof. Suppose $E = E_0 + i\Gamma$ in

$$\Psi = \psi e^{-iEt/\hbar} \quad (1.2.18)$$

The normalisability criterion requires that

$$1 = \int_{-\infty}^{+\infty} \Psi^* \Psi dx \quad (1.2.19a)$$

$$= \int_{-\infty}^{+\infty} \psi^2 e^{(iE^* - iE)t/\hbar} dx \quad (1.2.19b)$$

which means that

$$iE^* - iE = 0 \quad (1.2.20a)$$

$$i(E_0 - i\Gamma) - i(E_0 + i\Gamma) = 0 \quad (1.2.20b)$$

$$2\Gamma = 0 \quad (1.2.20c)$$

$$\Gamma = 0. \quad (1.2.20d)$$

□

Theorem 1.2.1.3 (ψ can always be taken to be real).

The time-independent wavefunction ψ , even if complex, can always be expressed as a linear combination of solutions with the same energy that are real (with complex coefficients).

Proof. For any ψ that satisfies

$$H\psi = E\psi, \quad (1.2.21)$$

ψ^* will also satisfy it. Therefore a linear combination of the two, *i.e.*, the real and the imaginary parts, will also separately satisfy the equation:

$$H(\psi + \psi^*) = E(\psi + \psi^*) \quad (1.2.22)$$

$$Hi(\psi - \psi^*) = Ei(\psi - \psi^*).$$

Therefore, we can always take ψ 's to be real when working with them. □

Theorem 1.2.1.4 (Even potentials).

If $V(x)$ is an even function, then ψ can always be taken to be either even or odd.

Proof. For any $\psi(x)$ that satisfies

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x) \quad (1.2.23)$$

$$H\psi(x) = E\psi(x)$$

with $V(x) = V(-x)$. $\psi(-x)$ can also satisfy the same equation:

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(-x)}{dx^2} + V(x)\psi(-x) = E\psi(-x) \quad (1.2.24)$$

$$H\psi(-x) = E\psi(-x),$$

Therefore in the same vein as the previous proof, two linear combinations $\psi_{even} = \psi(x) + \psi(-x)$ and $\psi_{odd} = \psi(x) - \psi(-x)$ can both satisfy the equation. \square

Theorem 1.2.1.5 (E must exceed minimum value of V).

E must exceed minimum value of V for every normalisable solution to the time-independent Schrödinger equation.

Proof. The time-independent Schrödinger equation can be rewritten as

$$\frac{d^2\psi}{dx^2} = \frac{2m}{\hbar^2}(V - E)\psi. \quad (1.2.25)$$

If $E < V_{min}$, ψ and ψ_{xx} will have the same sign, therefore maxima must occur when $\psi < 0$ and minima when $\psi > 0$, i.e., it always curves away from the x -axis, hence it will not tend to 0 as $|x| \rightarrow \infty$. The classical analogue is that $T \equiv K + V \geq V$. \square

1.2.2 The infinite square well

We now specify a potential:

$$V(x) = \begin{cases} 0 & \text{if } 0 \leq x \leq a, \\ \infty & \text{otherwise.} \end{cases} \quad (1.2.26)$$

We first solve the time-independent Schrödinger equation :

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} = E\psi$$

$$\frac{d^2\psi}{dx^2} = -k^2\psi, \text{ where } k \equiv \frac{\sqrt{2mE}}{\hbar}. \quad (1.2.27)$$

We arrive at the general SHM solution of

$$\psi(x) = A \sin kx + B \cos kx. \quad (1.2.28)$$

We impose the boundary condition

$$\psi(0) = \psi(a) = 0 \quad (1.2.29)$$

to yield

$$B = 0 \text{ and } k_n = \frac{n\pi}{a}. \quad (1.2.30)$$

Therefore from Equation 1.2.27 we have permitted values of E :

$$E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{n^2 \pi^2 \hbar^2}{2ma^2}. \quad (1.2.31)$$

We then find A_n by normalising ψ :

$$\int_0^a |A|^2 \sin^2(kx) dx = |A|^2 \cdot \frac{2}{a} = 1, \quad A = \sqrt{\frac{2}{a}}. \quad (1.2.32)$$

The second equality comes from the fact that $k = n\pi/a$, and the third is due to that the phase of A bears no physical significance. We now have

$$\psi_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right), \text{ and } E_n = \frac{n^2 \pi^2 \hbar^2}{2ma^2}. \quad (1.2.33)$$

Property 1.2.2.1 (Particle-in-a-box solutions). The solutions have the following properties:

1. They are *alternately even and odd* wrt the centre of the well: ψ_1 is even, ψ_2 is odd and so on.
2. Going up in energy, each successive state has one more node with ψ_1 having none.
3. They are *mutually orthogonal*:

$$\int_{-\infty}^{+\infty} \psi_m^* \psi_n dx = \delta_{mn}. \quad (1.2.34)$$

4. They are *complete* in that they are terms of an (odd) Fourier series and as such their linear combination can represent any (odd) function with the same period. The coefficients can be evaluated as follows

$$c_n = \int_{-\infty}^{+\infty} \psi_n^*(x) f(x) dx. \quad (1.2.35)$$

The most general solution is then

$$\Psi(x, t) = \sum_{n=1}^{\infty} c_n \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right) e^{-i(n^2\pi^2\hbar^2/2ma^2)t}, \quad (1.2.36)$$

where c_n can be appropriately chosen to fit any initial wavefunction $\Psi(x, 0)$:

$$c_n = \sqrt{\frac{2}{a}} \int_0^a \sin\left(\frac{n\pi}{a}x\right) \Psi(x, 0) dx. \quad (1.2.37)$$

Worked example 1.2.2.1. A particle in the infinite square well has the initial wavefunction

$$\Psi(x, 0) = Ax(a - x). \quad (0 \leq x \leq a) \quad (1.2.38)$$

Find $\Psi(x, t)$.

Solution: We normalise the boundary conditions first:

$$\begin{aligned} A &= \frac{1}{\sqrt{\int_0^a x^2(a - x)^2 dx}} \\ &= \sqrt{\frac{30}{a^5}}. \end{aligned} \quad (1.2.39)$$

The Fourier coefficients are

$$c_n = \frac{2\sqrt{15}}{a^3} \int_0^a \sin\left(\frac{n\pi}{a}x\right) x(a - x) dx \quad (1.2.40a)$$

$$= \frac{2\sqrt{15}}{a^2} \int_0^a x \sin\left(\frac{n\pi}{a}x\right) dx - \frac{2\sqrt{15}}{a^3} \int_0^a x^2 \sin\left(\frac{n\pi}{a}x\right) dx \quad (1.2.40b)$$

$$= \frac{2\sqrt{15}}{a^2} \left[\frac{-a^2}{n\pi} (-1)^n \right] - \frac{2\sqrt{15}}{a^3} \left[\frac{-a^3}{n\pi} (-1)^n - \frac{2a^3}{n^3\pi^3} [(-1)^n - 1] \right] \quad (1.2.40c)$$

$$= \frac{4\sqrt{15}}{(n\pi)^3} [(-1)^n - 1], \quad (1.2.40d)$$

which is equal to

$$\begin{cases} 0, & \text{if } n \text{ is even} \\ 8\sqrt{15}/(n\pi)^3, & \text{if } n \text{ is odd} \end{cases} \quad (1.2.41)$$

Therefore the complete wavefunction is

$$\Psi(x, t) = \sqrt{\frac{30}{a}} \left(\frac{2}{\pi}\right)^3 \sum_{\text{odd } n} \frac{1}{n^3} \sin\left(\frac{n\pi}{a}x\right) e^{-iEt}. \quad (1.2.42)$$

Theorem 1.2.2.1 (Coefficient squared sums to 1).

$$\sum_{n=1}^{\infty} |c_n|^2 = 1. \quad (1.2.43)$$

Proof.

$$1 = \int_{-\infty}^{+\infty} |\Psi(x, 0)|^2 dx \quad (1.2.44a)$$

$$= \int_{-\infty}^{+\infty} \left(\sum_{m=1}^{\infty} c_m \psi_m(x) \right)^* \left(\sum_{n=1}^{\infty} c_n \psi_n(x) \right) dx \quad (1.2.44b)$$

$$= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} c_m^* c_n \delta_{mn} \quad (1.2.44c)$$

$$= \sum_{n=1}^{\infty} |c_n|^2 \quad (1.2.44d)$$

□

Theorem 1.2.2.2 (Energy expectation in terms of coefficients).

$$\langle H \rangle = \sum_{n=1}^{\infty} |c_n|^2 E_n. \quad (1.2.45)$$

Proof. We know that

$$H\psi_n = E_n\psi_n. \quad (1.2.46)$$

Therefore we can write

$$\langle H \rangle = \int_{-\infty}^{+\infty} \Psi^* H \Psi dx \quad (1.2.47a)$$

$$= \int_{-\infty}^{+\infty} \left(\sum c_m \psi_m \right)^* H \left(\sum c_n \psi_n \right) dx \quad (1.2.47b)$$

$$= \sum \sum c_m^* c_n E_n \int_{-\infty}^{+\infty} \psi_m^* \psi_n dx = \sum |c_n|^2 E_n. \quad (1.2.47c)$$

□

Worked example 1.2.2.2. Calculate $\langle x \rangle$, $\langle x^2 \rangle$, $\langle p \rangle$, $\langle p^2 \rangle$, σ_x and σ_p .

A tedious exercise in integration, we will simply list the results

- $\langle x \rangle = a/2$

- $\langle x^2 \rangle = a^2(1/3 - 1/2n^2\pi^2)$
- $\langle p \rangle = 0$
- $\langle p^2 \rangle = n^2\pi^2\hbar^2/a^2$
- $\sigma_x^2 = (n^2\pi^2 - 6)a^2/12n^2\pi^2$
- $\sigma_p^2 = n^2\pi^2\hbar^2/a^2$
- $\sigma_x\sigma_p = \hbar\sqrt{(\pi^2n^2 - 6)/12}$
- Smallest uncertainty occurs when $n = 1$ at $\approx 0.568\hbar$.

Worked example 1.2.2.3. A particle in the infinite square well has as its initial wavefunction an even mixture of the first two stationary states:

$$\Psi(x, 0) = A[\psi_1(x) + \psi_2(x)]. \quad (1.2.48)$$

(a) Normalise $\Psi(x, 0)$.

$$1 = \int_0^a \Psi^* \Psi \, dx \quad (1.2.49a)$$

$$= \int_0^a A^2(\psi_1^2 + \psi_2^2 + 2\psi_1\psi_2) \, dx \quad (1.2.49b)$$

$$= 2A^2, \quad (1.2.49c)$$

which implies $A = 1/\sqrt{2}$. We invoke Theorem 1.1.1.1 to be assured that we only need to normalise the wavefunction this one time.

(b) Find $\Psi(x, t)$ and $|\Psi(x, t)|^2$.

By properties of Fourier series we arrive immediately at

$$\begin{aligned} \Psi(x, t) &= \frac{1}{\sqrt{2}} [\psi_1 e^{-i\omega t} + \psi_2 e^{-4i\omega t}] \\ &= \frac{\sqrt{a}}{2} [\sin(\pi x/a) e^{-i\omega t} + \sin(2\pi x/a) e^{-4i\omega t}], \end{aligned} \quad (1.2.50)$$

where we have set $\omega \equiv \pi^2\hbar/2ma^2$. Then

$$|\Psi(x, t)|^2 = \frac{1}{2} [\psi_1^2 + \psi_2^2 + 2\psi_1\psi_2 \cos(3\omega t)] \quad (1.2.51)$$

(c) Compute $\langle x \rangle$ and $\langle p \rangle$.

The integral to evaluate is

$$\begin{aligned} & \frac{1}{a} \int_0^a x \sin^2(\pi x/a) + x \sin^2(2\pi x/a) + 2x \sin(\pi x/a) \sin(2\pi x/a) \cos(3\omega t) dx \\ &= \frac{1}{a} \int_0^a x \sin^2(\pi x/a) + x \sin^2(2\pi x/a) + 2x \sin^2(\pi x/a) \cos(\pi x/a) \cos(3\omega t) dx \end{aligned} \quad (1.2.52)$$

Tedious again, we evaluate the integrals separately, using double angle and trigo power expansion formulae in the second integral:

$$\int_0^a x \sin\left(\frac{n\pi}{a}\right) dx = \frac{a^2}{4} \quad (1.2.53a)$$

$$\int_0^a 2x \sin^2\left(\frac{\pi x}{a}\right) \cos\left(\frac{\pi x}{a}\right) \cos(3\omega t) dx = -\frac{8a}{9\pi^2} \cos(3\omega t) \quad (1.2.53b)$$

So,

$$\langle x \rangle = \frac{2}{a} - \frac{16a}{9\pi^2} \cos(3\omega t). \quad (1.2.54)$$

$\langle p \rangle$ can easily be evaluated by Ehrenfest theorem (eq. 1.1.14a):

$$\langle p \rangle = \frac{16m\omega a}{3\pi^2} \sin(3\omega t) = \frac{8\hbar}{3a} \sin(3\omega t). \quad (1.2.55)$$

(d) If you measured the energy of this particle, what values might you get and what's the probability of getting each of them? Find $\langle H \rangle$.

We might get $E_1 = \pi^2 \hbar^2 / 2ma^2$ and $E_2 = 2\pi^2 \hbar^2 / ma^2$ each with 0.5 probability, and $\langle H \rangle = 1.25\pi^2 \hbar^2 / ma^2$.

Worked example 1.2.2.4. A oarticle in the infinite square well has the initial wavefunction

$$\Psi(x, 0) = \begin{cases} Ax, & 0 \leq x \leq a/2 \\ A(a-x), & a/2 \leq x \leq a \end{cases} \quad (1.2.56)$$

Find $\Psi(x, t)$

Solution: We normalise the initial wavefunction first:

$$1 = \int_0^a \Psi^* \Psi dx \quad (1.2.57a)$$

$$= \frac{a^3 A^2}{24} + \frac{a^3 A^2}{24} = \frac{a^3 A^2}{12}, \quad (1.2.57b)$$

which means that $A = \sqrt{12}/a^{3/2}$. The coefficients are obtained by

$$c_n = \int_0^{a/2} Ax \cdot \sqrt{\frac{2}{a}} \sin(n\pi x/a) dx + \int_{a/2}^a A(a-x) \sqrt{\frac{2}{a}} \sin(n\pi x/a) dx \quad (1.2.58a)$$

$$= \frac{2\sqrt{6}}{a^2} \left[\int_0^{a/2} x \sin(n\pi x/a) dx - \int_{a/2}^a x \sin(n\pi x/a) dx + \int_{a/2}^a a \sin(n\pi x/a) dx \right] \quad (1.2.58b)$$

$$= \frac{2\sqrt{6}}{a^2} \left[\frac{a^2}{n^2\pi^2} \sin(n\pi/2) + \frac{a^2}{n\pi} \cos(n\pi/2) + \frac{a^2}{n^2\pi^2} \sin(n\pi/2) - \frac{a^2}{n\pi} \cos(n\pi/2) \right] \quad (1.2.58c)$$

$$= \frac{4\sqrt{6}}{n^2\pi^2} \sin(n\pi/2). \quad (1.2.58d)$$

We write the initial conditions as a linear combination of the stationary states:

$$\Psi(x, 0) = \frac{4\sqrt{6}}{\pi^2} \left[\sum_{n=1} \frac{\psi_n}{(4n-3)^2} - \frac{\psi_n}{(4n-1)^2} \right]. \quad (1.2.59)$$

Full solution can be written out:

$$\Psi(x, t) = \frac{4\sqrt{6}}{\pi^2} \left[\sum_{n=1} \frac{\psi_n}{(4n-3)^2} e^{-iE_{4n-3}t/\hbar} - \frac{\psi_n}{(4n-1)^2} e^{-iE_{4n-1}t/\hbar} \right]. \quad (1.2.60)$$

The probability of energy turning out to be E_1 is

$$|c_1|^2 = \frac{96}{\pi^4} \approx 0.9855. \quad (1.2.61)$$

And

$$\langle H \rangle = \sum_{n=1} |c_n|^2 E_n = \frac{48\hbar^2}{2\pi^2 ma^2} \sum_{\text{odd } n} \frac{1}{n^2} = \frac{48\hbar^2}{2\pi^2 ma^2} \frac{\pi^2}{8} = \frac{6\hbar^2}{ma^2} \approx 1.216 E_1. \quad (1.2.62)$$

This example illustrates that the initial wavefunction does *not* need to have continuous first or second derivative or, for that matter, be continuous.

1.2.3 The harmonic oscillator

The harmonic potential is a very important class of potentials due to the ubiquitous presence of it as a result of Taylor expansion of many more complex potentials. It is given as

$$V(x) = \frac{1}{2}m\omega^2 x^2, \quad (1.2.63)$$

where we can replace m with μ when we are dealing with heteronuclear diatomics. The corresponding time-independent Schrödinger equation is

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + \frac{1}{2}m\omega^2 x^2 \psi = E\psi. \quad (1.2.64)$$

Two methods at solving are available, we discuss the algebraic 'ladder operator' solution first.

1.2.3.1 Ladder operator

We can rewrite Equation 1.2.64 in a more suggestive form

$$\frac{1}{2m}[p^2 + (m\omega x)^2]\psi = E\psi, \quad (1.2.65)$$

where we have rewritten the Hamiltonian operator in terms of momentum and position operators.

We are therefore inspired to decompose it into two operators of the form

$$q_{\pm} \equiv \mp ip + m\omega x. \quad (1.2.66)$$

Aside: we do not write $q_{\pm} \equiv p \pm im\omega x$ because we want to obtain real ψ 's (Theorem 1.2.1.3) and the i in front of p operator does that for us by getting rid of the i in the momentum operator.

The two operators should ideally be such that

$$q_+ q_- \stackrel{?}{=} p^2 + (m\omega x)^2. \quad (1.2.67)$$

But we cannot do this as p and x commutes, and we should calculate the commutator $[x, p]$ first.

$$[x, p]f(x) = \left[x \frac{\hbar}{i} \frac{d}{dx}(f) - \frac{\hbar}{i} \frac{d}{dx}(xf) \right] \quad (1.2.68a)$$

$$= \frac{\hbar}{i} \left(x \frac{df}{dx} - f - x \frac{df}{dx} \right) \quad (1.2.68b)$$

$$= i\hbar f(x) \quad (1.2.68c)$$

Property 1.2.3.1. This gives the **canonical commutation relation**

$$[x, p] = i\hbar \quad (1.2.68d)$$

We re-examine q_+q_- :

$$q_+q_- = (-ip + m\omega x)(ip + m\omega x) = p^2 + (m\omega x)^2 + im\omega[x, p] = 2mH - m\omega\hbar. \quad (1.2.68e)$$

This means that

$$H \equiv \hbar\omega \left(\frac{q_+q_-}{2m\hbar\omega} + \frac{1}{2} \right) \quad (1.2.69)$$

For simplicity's sake we redefine

Definition 1.2.3.1 (Ladder operators).

$$a_{\pm} \equiv \frac{1}{\sqrt{2m\hbar\omega}}(\mp ip + m\omega x), \quad (1.2.70)$$

So now we have

$$H = \hbar\omega \left(a_+a_- + \frac{1}{2} \right) = \hbar\omega \left(a_-a_+ - \frac{1}{2} \right) = \hbar\omega \left(a_{\pm}a_{\mp} \pm \frac{1}{2} \right). \quad (1.2.71)$$

Property 1.2.3.2 (Ladder operator commutator). Additionally, the commutator

$$[a_-, a_+] = 1 \quad (1.2.72)$$

will come in handy.

We now introduce the important theorem

Theorem 1.2.3.1 (Ladder operator theorem). If ψ satisfies the Schrödinger equation with energy E then $a_+\psi$ satisfies the Schrödinger equation with energy $E + \hbar\omega$.

Proof.

$$H(a_+\psi) = \hbar\omega \left(a_+a_- + \frac{1}{2} \right) (a_+\psi) = \hbar\omega \left(a_+a_-a_+ + \frac{1}{2}a_+ \right) \psi \quad (1.2.73a)$$

$$= \hbar\omega a_+ \left(a_-a_+ + \frac{1}{2} \right) \psi = a_+ \left[\hbar\omega \left(a_+a_- + 1 + \frac{1}{2} \right) \psi \right] \quad (1.2.73b)$$

$$= a_+(H + \hbar\omega)\psi = a_+(E + \hbar\omega)\psi = (E + \hbar\omega)(a_+\psi). \quad (1.2.73c)$$

Similarly,

$$H(a_-\psi) = (E - \hbar\omega)(a_-\psi). \quad (1.2.74)$$

□

Corollary 1.2.3.1 (Ground state of harmonic potential). The normalised ground state is

$$\psi_0 = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-m\omega x^2/2\hbar}. \quad (1.2.75)$$

with energy

$$E_0 = \frac{1}{2}\hbar\omega. \quad (1.2.76)$$

Proof. The ground state occurs when the lowering operator fails to produce a normalisable wavefunction, i.e.,

$$0 = a_-\psi_0 \quad (1.2.77a)$$

$$= \frac{1}{\sqrt{2\hbar m\omega}} \left(\hbar \frac{d}{dx} + m\omega x \right) \psi_0, \quad (1.2.77b)$$

or,

$$\frac{d\psi_0}{dx} = -\frac{m\omega}{\hbar} x \psi_0. \quad (1.2.78)$$

Solutions is obtained immediately as

$$\psi_0 = A e^{-\frac{m\omega}{2\hbar} x^2}. \quad (1.2.79)$$

To normalise,

$$A = \frac{1}{\sqrt{\int_{-\infty}^{+\infty} e^{-\frac{m\omega}{\hbar} x^2}}} = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4}. \quad (1.2.80)$$

Therefore the ground state wavefunction is

$$\psi_0 = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-m\omega x^2/2\hbar}. \quad (1.2.81)$$

The energy of the ground state can be found as follows

$$\hbar\omega \left(a_+ a_- + \frac{1}{2} \right) \psi_0 = E_0 \psi_0 \quad (1.2.82a)$$

$$\frac{1}{2} \hbar\omega \psi_0 + \hbar\omega a_+ a_- \psi_0 = E_0 \psi_0 \quad (1.2.82b)$$

$$E_0 = \frac{1}{2} \hbar\omega. \quad (1.2.82c)$$

□

We can therefore now write¹

$$\psi_n(x) = A_n(a_+)^n \psi_0(x), \text{ with } E_n = \left(n + \frac{1}{2}\right) \hbar\omega, \quad (1.2.83)$$

where A_n is the normalisation factor (on top of the normalisation factor that is included in the ground state wavefunction), and we will determine it in Theorem 1.2.3.2. But to do so we first need to introduce a lemma:

Lemma 1.2.3.1 (Hermitian conjugate). For any well-behaved, *i.e.*, goes to zero at $\pm\infty$, functions $f(x)$ and $g(x)$, we have, where a_{\pm} and a_{\mp} can be any pair of hermitian conjugates,

$$\int_{-\infty}^{+\infty} f^*(a_{\pm}g) dx = \int_{-\infty}^{+\infty} (a_{\mp}f)^* g dx. \quad (1.2.84)$$

Proof. We show this result by explicit evaluation of the integral:

$$\int_{-\infty}^{+\infty} f^*(a_{\pm}g) dx = \frac{1}{\sqrt{2\hbar m\omega}} \int_{-\infty}^{+\infty} f^* \left(\mp \hbar \frac{d}{dx} + m\omega x \right) g dx \quad (1.2.85a)$$

$$= \frac{1}{\sqrt{2\hbar m\omega}} \int_{-\infty}^{+\infty} \left[\left(\pm \hbar \frac{d}{dx} + m\omega x \right) f \right]^* g dx \quad (1.2.85b)$$

$$= \int_{-\infty}^{+\infty} (a_{\mp}f)^* g dx. \quad (1.2.85c)$$

Where the second last equality comes from integration by parts with the boundary terms vanishing for well-behaved functions. □

Theorem 1.2.3.2 (Normalisation of harmonic potential wavefunctions).

$$\psi_n = \frac{1}{\sqrt{n!}} (a_+)^n \psi_0. \quad (1.2.86)$$

Proof. We know that $a_{\pm}\psi_n$ is proportional to $\psi_{n\pm 1}$, with all three of the wavefunctions normalised, so we have

$$a_+\psi_n = c_n\psi_{n+1}, \quad a_-\psi_n = d_n\psi_{n-1}. \quad (1.2.87)$$

¹Many sources prefer the use of v as the index here to distinguish between the vibrational quantum number and the principal quantum number of hydrogen atom. We will adopt this notation in later chapters.

From Lemma 1.2.3.1 we can write that

$$\int_{-\infty}^{+\infty} (a_{\pm}\psi_n)^*(a_{\pm}\psi_n) dx = \int_{-\infty}^{+\infty} (a_{\mp}a_{\pm}\psi_n)^*\psi_n dx. \quad (1.2.88)$$

We then invoke Equations 1.2.71 and 1.2.83 and write

$$H\psi_n = E_n\psi_n \quad (1.2.89a)$$

$$\hbar\omega \left(a_+a_- + \frac{1}{2} \right) \psi_n = \hbar\omega \left(n + \frac{1}{2} \right) \psi_n \quad (1.2.89b)$$

$$a_+a_-\psi_n = n\psi_n. \quad (1.2.89c)$$

The commutator relation further gives us that

$$a_-a_+\psi_n = (n+1)\psi_n. \quad (1.2.90)$$

So it follows that

$$\int_{-\infty}^{+\infty} (a_+\psi_n)^*(a_+\psi_n) dx = |c_n|^2 \int_{-\infty}^{+\infty} |\psi_{n+1}|^2 dx = (n+1) \int_{-\infty}^{+\infty} |\psi_n|^2 dx \quad (1.2.91a)$$

$$\int_{-\infty}^{+\infty} (a_-\psi_n)^*(a_-\psi_n) dx = |d_n|^2 \int_{-\infty}^{+\infty} |\psi_{n-1}|^2 dx = n \int_{-\infty}^{+\infty} |\psi_n|^2 dx. \quad (1.2.91b)$$

Because we know that all three wavefunctions are normalised, it must follow that $|c_n|^2 = n+1$ and that $|d_n|^2 = n$, and hence we obtain the recurrence relation

$$a_+\psi_n = \sqrt{n+1}\psi_{n+1}, \quad a_-\psi_n = \sqrt{n}\psi_{n-1}. \quad (1.2.92)$$

Finally,

$$\psi_n = \frac{1}{\sqrt{n!}}(a_+)^n\psi_0. \quad (1.2.93)$$

□

Theorem 1.2.3.3 (Orthonormality). Stationary states of the harmonic oscillator are orthogonal:

$$\int_{-\infty}^{+\infty} \psi_m^*\psi_n dx = \delta_{mn}. \quad (1.2.94)$$

Proof.

$$\int_{-\infty}^{+\infty} \psi_m^* (a_+ a_-) \psi_n dx = n \int_{-\infty}^{+\infty} \psi_m^* \psi_n dx \quad (1.2.95a)$$

$$= \int_{-\infty}^{+\infty} (a_- \psi_m)^* (a_- \psi_n) dx \quad (1.2.95b)$$

$$= \int_{-\infty}^{+\infty} (a_+ a_- \psi_m)^* \psi_n dx \quad (1.2.95c)$$

$$= m \int_{-\infty}^{+\infty} \psi_m^* \psi_n dx, \quad (1.2.95d)$$

where the first equality comes from Equation 1.2.89c, and the second and third from Equation 1.2.84. Unless $m = n$ the integral must vanish.

Orthonormality means that we can again use the Fourier series method to evaluate the coefficients of $\Psi(x, 0)$ when it is written as a linear combination of stationary states, and $|c_n|^2$ is again the probability that a measurement of the energy would yield the value E_n . \square

Worked example 1.2.3.1. Find the expectation value of the potential energy in the n -th state of the harmonic oscillator.

Solution:

We need to evaluate

$$\langle V \rangle = \left\langle \frac{1}{2} m \omega^2 x^2 \right\rangle = \frac{1}{2} m \omega^2 \int_{-\infty}^{+\infty} \psi_n^* x^2 \psi_n dx. \quad (1.2.96)$$

The ladder operators again comes in very handy as we can use them to evaluate integrals involving powers of x and p :

$$x = \sqrt{\frac{\hbar}{2m\omega}} (a_+ + a_-) \quad (1.2.97a)$$

$$p = i \sqrt{\frac{\hbar m \omega}{2}} (a_+ - a_-). \quad (1.2.97b)$$

So

$$x^2 = \frac{\hbar}{2m\omega} \left[(a_+)^2 + (a_+ a_-) + (a_- a_+) + (a_-)^2 \right], \quad (1.2.98)$$

and so

$$\langle V \rangle = \frac{\hbar \omega}{4} \int_{-\infty}^{+\infty} \psi_n^* \left[(a_+)^2 + (a_+ a_-) + (a_- a_+) + (a_-)^2 \right] \psi_n dx. \quad (1.2.99)$$

The squares of raising and lowering operators will result in $\psi_{n\pm 2}$ and as such will drop out due to the orthonormality condition. The middle two terms can be evaluated by Equa-

tions 1.2.90 and 1.2.89c to get

$$\langle V \rangle = \frac{\hbar\omega}{4}(n + n + 1) = \frac{1}{2}\hbar\omega \left(n + \frac{1}{2} \right), \quad (1.2.100)$$

exactly half of $\langle E \rangle$.

Worked example 1.2.3.2. Construct $\psi_2(x)$.

Solution:

We just have to apply the raising operator twice:

We have

$$\psi_0 = \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} e^{-m\omega x^2/2\hbar}. \quad (1.2.101)$$

So,

$$\psi_1 = \frac{1}{\sqrt{1!}} \frac{1}{\sqrt{2m\hbar\omega}} \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} \left(-\hbar \frac{d}{dx} + m\omega x \right) e^{-m\omega x^2/2\hbar} \quad (1.2.102a)$$

$$= \frac{2m\omega}{\sqrt{2m\hbar\omega}} \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} x e^{-m\omega x^2/2\hbar} \quad (1.2.102b)$$

$$= \frac{\sqrt{2}}{\pi^{1/4}} \left(\frac{m\omega}{\hbar} \right)^{3/4} x e^{-m\omega x^2/2\hbar}. \quad (1.2.102c)$$

Applying the raising operator again (ugh) we get

$$\psi_2 = \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2m\hbar\omega}} \left(\hbar \frac{d}{dx} + m\omega x \right) x e^{-m\omega x^2/2\hbar} \quad (1.2.103a)$$

$$= \left(\frac{m\omega}{4\pi\hbar} \right)^{1/4} \left(2\frac{m\omega}{\hbar} x^2 - 1 \right) e^{-m\omega x^2/2\hbar}. \quad (1.2.103b)$$

Property 1.2.3.3 (Symmetry). The stationary states of the harmonic oscillator are alternately even and odd, with ψ_0 even.

Property 1.2.3.4 (Substitutions). We can introduce the following substitutions to simplify calculations. We will see that these substitutions also arise from the power series solution discussed below.

$$\xi \equiv \sqrt{\frac{m\omega}{\hbar}} x \quad (1.2.104a)$$

$$\alpha \equiv \left(\frac{m\omega}{\pi\hbar} \right)^{1/4}. \quad (1.2.104b)$$

For example, ψ_0 now becomes

$$\psi_0 = \alpha e^{-\xi^2/2}. \quad (1.2.105)$$

Property 1.2.3.5 (Integrals containing exponentials). We give some common integrals, all of the integrals involving x^n , where n is a positive integer, can be proven by induction:

$$\int_0^\infty x^n e^{-ax} dx = \frac{n!}{a^{n+1}} \quad (1.2.106a)$$

$$\int_0^\infty e^{-ax^2} dx = \left(\frac{\pi}{4a}\right)^{1/2} \quad (1.2.106b)$$

$$\int_0^\infty x^{2n} e^{-ax^2} dx = \frac{\prod_{\text{odd } r, r=1}^{2n-1} (r)}{(2a)^n} \left(\frac{\pi}{4a}\right)^{1/2} \quad (1.2.106c)$$

$$\int_0^\infty x^{2n+1} e^{-ax^2} dx = \frac{n!}{2a^{n+1}}. \quad (1.2.106d)$$

Worked example 1.2.3.3. Find $\langle x \rangle$, $\langle p \rangle$, $\langle x^2 \rangle$, $\langle p^2 \rangle$ and $\langle T \rangle$ for ψ_0 and ψ_1 by explicit integration.

Solution:

For ψ_0 , an even function, we have

$$\langle x \rangle = \langle p \rangle = 0. \quad (1.2.107)$$

And

$$\langle x^2 \rangle = \int_{-\infty}^{+\infty} \psi_0^* x^2 \psi_0 dx \quad (1.2.108a)$$

$$= \alpha^2 \left(\frac{\hbar}{m\omega}\right)^{\frac{3}{2}} \int_{-\infty}^{+\infty} \xi^2 e^{-\xi^2} d\xi \quad (1.2.108b)$$

$$= \alpha^2 \left(\frac{\hbar}{m\omega}\right)^{\frac{3}{2}} \frac{\sqrt{\pi}}{2} \quad (1.2.108c)$$

$$= \frac{\hbar}{2m\omega}, \quad (1.2.108d)$$

where we used the standard result that $\int_{-\infty}^{+\infty} \xi^2 e^{-\xi^2} d\xi = \sqrt{\pi}/2$. For momentum, we first note that the momentum operator under our substitutions is

$$p = -i\hbar \sqrt{\frac{m\omega}{\hbar}} \frac{d}{d\xi} \quad (1.2.109a)$$

$$p^2 = -m\omega\hbar \frac{d^2}{d\xi^2}. \quad (1.2.109b)$$

We therefore have

$$\langle p^2 \rangle = -m\omega\hbar\alpha^2 \sqrt{\frac{\hbar}{m\omega}} \int_{-\infty}^{+\infty} e^{-\xi^2/2} \frac{d^2}{d\xi^2} e^{-\xi^2/2} d\xi \quad (1.2.110a)$$

$$= m\omega\hbar\pi^{-\frac{1}{2}} \int_{-\infty}^{+\infty} (\xi^2 e^{-\xi^2} - e^{-\xi^2}) d\xi \quad (1.2.110b)$$

$$= \frac{1}{2}m\omega\hbar. \quad (1.2.110c)$$

Therefore the uncertainty principle for ψ_0 is

$$\sigma_x \sigma_p = \frac{\hbar}{2}, \quad (1.2.111)$$

the smallest possible uncertainty. Now, for ψ_1 , an odd function, we immediately know that $x\psi_1^2$ is odd. The derivative of an odd function is even, and as such $\psi_1 p \psi_1$ is odd, and so

$$\langle x \rangle = \langle p \rangle = 0. \quad (1.2.112)$$

Moving on to the squares, we have the wavefunction

$$\psi_1 = \sqrt{\frac{2\pi\hbar}{m\omega}} \alpha^3 \xi e^{-\xi^2/2}. \quad (1.2.113)$$

So,

$$\langle x^2 \rangle = 2\pi \left(\frac{\hbar}{m\omega} \right)^{5/2} \alpha^6 \int_{-\infty}^{+\infty} \xi^4 e^{-\xi^2} d\xi \quad (1.2.114a)$$

$$= 2\pi \left(\frac{\hbar}{m\omega} \right)^{5/2} \alpha^6 \frac{1 \cdot 3}{2^2} \sqrt{\pi} \quad (1.2.114b)$$

$$= \frac{3}{2} \frac{\hbar}{m\omega}. \quad (1.2.114c)$$

And,

$$\langle p^2 \rangle = -2\pi\hbar^2 \left(\frac{\hbar}{m\omega} \right)^{3/2} \alpha^6 \int_{-\infty}^{+\infty} (\xi^4 - 3\xi^2) e^{-\xi^2} d\xi \quad (1.2.115a)$$

$$= \frac{3}{2} m\omega\hbar. \quad (1.2.115b)$$

And therefore the uncertainty principle is

$$\sigma_x \sigma_p = \frac{3\hbar}{2}. \quad (1.2.116)$$

Finally we note that

$$\langle T \rangle = \frac{\langle p^2 \rangle}{2m} \quad (1.2.117a)$$

$$\langle V \rangle = \frac{k \langle x^2 \rangle}{2} \quad (1.2.117b)$$

So we have

$$\langle T \rangle_0 = \langle V \rangle_0 = \frac{\hbar\omega}{4} \quad (1.2.118a)$$

$$\langle T \rangle_1 = \langle V \rangle_1 = \frac{3\hbar\omega}{4}, \quad (1.2.118b)$$

as expected.

Worked example 1.2.3.4. Now calculate the same quantities for ψ_n .

Solution:

We use the method developed in Worked example 1.2.3.1 (Equations 1.2.97a and 1.2.97b).

As x and p only contains a_{\pm} terms, by orthonormality conditions,

$$\langle x \rangle_n = \langle p \rangle_n = 0. \quad (1.2.119)$$

And

$$\langle x^2 \rangle_n = \frac{\hbar}{2m\omega} \int_{-\infty}^{+\infty} \psi_n^* [(a_+ a_-) + (a_- a_+)] \psi_n dx \quad (1.2.120a)$$

$$= \frac{\hbar}{m\omega} \left(n + \frac{1}{2} \right), \quad (1.2.120b)$$

$$\langle p^2 \rangle_n = \frac{-\hbar m\omega}{2} \int_{-\infty}^{+\infty} \psi_n^* [-(a_+ a_-) - (a_- a_+)] \psi_n dx \quad (1.2.121a)$$

$$= \hbar m\omega \left(n + \frac{1}{2} \right). \quad (1.2.121b)$$

The uncertainty principle is then

$$(\sigma_x)_n (\sigma_p)_n = \hbar \left(n + \frac{1}{2} \right). \quad (1.2.121c)$$

Energies are

$$\langle T \rangle = \langle V \rangle = \frac{1}{2} \hbar\omega \left(n + \frac{1}{2} \right). \quad (1.2.121d)$$

Another thing that ladder operator can do is to derive the selection rule for harmonic oscillators. This will be derived in Section 2.2.3.

1.2.3.2 Analytic Method

The Schrödinger equation for the harmonic oscillator is

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + \frac{1}{2}m\omega^2 x^2 \psi = E\psi. \quad (1.2.122)$$

We now want to solve it directly via the power series method. We first non-dimensionalise the equation [24], because doing so will

- simplify the form of the equation greatly;
- reveal characteristic quantities of the equation;
- reduce the number of coefficients which may depend on other variables, and hence simplify the process of solving the equation numerically, if so required.

Non-dimensionalisation

To nondimensionalise ψ , the unknown to be solved, we proceed as follows:

First we note that, since

$$1 = \int_{-\infty}^{+\infty} |\psi(x)|^2 dx, \quad (1.2.123)$$

we can conclude that $|\psi(x)|^2$ has the unit of inverse length, we must write it in terms of a dimensionless variable. To do this, we introduce substitution

$$\xi \equiv \frac{x}{x_c}, \quad (1.2.124)$$

where x_c is some characteristic length of the system. This gives us the dimensionless wavefunction:

$$\psi(x) = \psi(\xi x_c) = \psi(x(x_c)) \equiv \tilde{\psi}(\xi), \quad (1.2.125)$$

where $\tilde{\psi}$ is the non-dimensionalised wavefunction that takes as its argument ξ and outputs the same value of ψ at corresponding x , hence the defined equivalence sign.

Now the Schrödinger equation becomes

$$\left(-\frac{\hbar^2}{2m} \frac{1}{x_c^2} \frac{d^2}{d\xi^2} + \frac{1}{2}m\omega^2 x_c^2 \xi^2 \right) \tilde{\psi} = E\tilde{\psi} \quad (1.2.126a)$$

$$\Rightarrow \left(-\frac{d^2}{d\xi^2} + \frac{m^2\omega^2 x_c^4}{\hbar^2} \xi^2 \right) \tilde{\psi} = \frac{2mx_c^2 E}{\hbar^2} \tilde{\psi}. \quad (1.2.126b)$$

To make the coefficients of ξ^2 dimensionless as well, we set

$$\frac{m^2 \omega^2 x_c^4}{\hbar^2} = 1 \Rightarrow x_c = \sqrt{\frac{\hbar}{m\omega}}. \quad (1.2.127)$$

Non-dimensionalising E as well:

$$K \equiv \frac{2E}{\hbar\omega}, \quad (1.2.128)$$

we arrive at the fully dimensionless equation

$$\frac{d^2\psi}{d\xi^2} = (\xi^2 - K)\psi, \quad (1.2.129)$$

where we have written $\tilde{\psi}$ as ψ for simplicity of notation, but it is important that we remember to revert it back after solving the equation.

An almost identical alternative nondimensionalisation scheme uses $\hbar\omega$ (instead of $\hbar\omega/2$ of K , that's the only difference) as the natural unit of energy and arrives at

$$\left(-\frac{1}{2} \frac{d^2}{dq^2} + \frac{1}{2} q^2 \right) \psi = E\psi, \quad (1.2.130)$$

where q is the scaled position coordinate, and is identically defined as ξ .

Asymptotic analysis

At large $|\xi|$, ξ^2 dominates over the constant K , so in this regime

$$\frac{d^2\psi}{d\xi^2} \approx \xi^2\psi, \quad (1.2.131)$$

which gives

$$\psi(\xi) \approx Ae^{-\xi^2/2} + Be^{\xi^2/2}. \quad (1.2.132)$$

To get a normalisable wavefunction, we must have that $B = 0$.

Frobenius method

We try the solution of the form

$$\psi(\xi) = h(\xi)e^{-\xi^2/2}. \quad (1.2.133)$$

Substituting Equation 1.2.133 into Equation 1.2.129, we first differentiate it twice:

$$\frac{d^2\psi}{d\xi^2} = \left(\frac{d^2h}{d\xi^2} - 2\xi \frac{dh}{d\xi} + (\xi^2 - 1)h \right) e^{-\xi^2/2}, \quad (1.2.134)$$

which enables us to write

$$\frac{d^2 h}{d\xi^2} - 2\xi \frac{dh}{d\xi} + (K - 1)h = 0. \quad (1.2.135)$$

Performing a series expansion in ξ ,

$$h(\xi) = \sum_{j=0}^{\infty} a_j \xi^j, \quad (1.2.136)$$

plugging this into Equation 1.2.129 we get

$$0 = \sum_{j=0}^{\infty} j(j-1)a_j \xi^{j-2} - \sum_{j=0}^{\infty} 2ja_j \xi^j + \sum_{j=0}^{\infty} (K-1)a_j \xi^j \quad (1.2.137a)$$

$$= \sum_{j=0}^{\infty} [(j+2)(j+1)a_{j+2} - 2ja_j + (K-1)a_j] \xi^j \quad (1.2.137b)$$

$$\Rightarrow (j+1)(j+2)a_{j+2} - 2ja_j + (K-1)a_j = 0. \quad (1.2.137c)$$

We get

$$a_{j+2} = \frac{2j+1-K}{(j+1)(j+2)} a_j, \quad (1.2.138)$$

a *recursion formula* with two arbitrary constants, a_0 and a_1 , which are fixed by

$$a_0 = h(0) \quad (1.2.139a)$$

$$a_1 = h'(0). \quad (1.2.139b)$$

It is entirely equivalent to the Schrödinger equation we started with.

We write the complete solution as

$$h(\xi) = h_{\text{even}}(\xi) + h_{\text{odd}}(\xi), \quad (1.2.140)$$

where

$$h_{\text{even}}(\xi) \equiv a_0 + a_2 \xi^2 + a_4 \xi^4 + \dots \quad (1.2.141)$$

and likewise for $h_{\text{odd}}(\xi)$.

Truncation

Up till now we have not checked if our solution is normalisable, *i.e.*, physical. We inspect the behaviour of recursion formula for suspicious behaviour at large j , when the recursion formula becomes

$$a_{j+2} \approx \frac{2}{j} a_j, \quad (1.2.142)$$

which means

$$a_j = \frac{2}{j-2} \cdot \frac{2}{j-4} \cdot \frac{2}{j-6} \cdots \approx \frac{C}{(j/2)!}. \quad (1.2.143)$$

This yields at large ξ where higher powers dominate

$$h(\xi) \approx C \sum \frac{1}{(j/2)!} \xi^j. \quad (1.2.144)$$

Relabelling, we get

$$h(\xi) \approx C \sum \frac{1}{j!} \xi^{2j} = C e^{\xi^2}. \quad (1.2.145)$$

This is not good as

$$\psi(\xi) = h(\xi) e^{-\xi^2} \approx e^{\xi^2}, \quad (1.2.146)$$

and as such is not normalisable. However we note that we assumed large j to get here, so we should terminate the series before it can reach large j . This will truncate either the even or odd series and the other one must be *set* to 0 from the start. Say the highest $j \equiv n$ is now set, and we see from Equation 1.2.138 that

$$K = 2n + 1. \quad (1.2.147)$$

So energy must be

$$E = \left(n + \frac{1}{2}\right) \hbar \omega \quad (1.2.148)$$

for integer n 's. For allowed values of K , the recursion formula in Equation 1.2.138 becomes

$$a_{j+2} = \frac{-2(n-j)}{(j+1)(j+2)} a_j. \quad (1.2.149)$$

There are a few features to this formula:

- For even n we must fix a definite non-zero value for a_0 , for odd n , a_1 .
- For even n a_1 must be fixed at 0 because the n value is unable to terminate the odd series, and vice versa for odd n .
- For each value of n the power series coefficients are different and are subject to normalisation.

Normalisation

The polynomials generated by this recursion formula are named **Hermite polynomials**, $H(\xi)$, and conventionally the a_i 's are chosen such that the highest power of ξ is 2^n . We demonstrate the generation of a few such polynomials:

For $n = 0$ we must pick $a_1 = 0$ and so

$$H_0(\xi) = 1. \quad (1.2.150)$$

For $n = 1$ we set $a_0 = 0$ so

$$H_1(\xi) = 2\xi. \quad (1.2.151)$$

For $n = 2$, $a_1 = 0$, $a_2 = -2a_0$, therefore,

$$H_2(\xi) = 4\xi^2 - 2, \quad (1.2.152)$$

and so on. Now the wavefunction is

$$\psi_n(x) = A_n H_n(\xi) e^{-\xi^2/2}, \quad (1.2.153)$$

and we need to find out the normalisation constant. This process is by no means straightforward and very often omitted by textbooks. The following outline of proof is adapted from p.69-72 in [14], Section 2.7 of [21] and [9]:

The Hermite equation is obtained when $K = 2n + 1$ in Equation 1.2.135:

$$H_n'' - 2\xi H_n' + 2n H_n = 0. \quad (1.2.154)$$

We introduce the generating function for Hermite polynomials, $S(\xi, s)$:

$$S(\xi, s) = e^{\xi^2 - (s - \xi)^2} = e^{-s^2 + 2s\xi}. \quad (1.2.155)$$

We perform a Taylor series expansion on the second expression about $\xi = 0$:

$$F(\xi, s) \equiv \frac{S(\xi, s)}{e^{\xi^2}} = e^{-(s - \xi)^2} = e^{-(\xi - s)^2} \quad (1.2.156a)$$

$$= \sum_{n=0}^{\infty} (-1)^n \frac{s^n}{n!} (e^{-\xi^2})^{(n)}, \quad (1.2.156b)$$

and so we define

$$S(\xi, s) = \sum_{n=0}^{\infty} H_n \xi \frac{s^n}{n!}, \quad (1.2.157)$$

with

$$H_n(\xi) = \frac{(-1)^n}{n!} e^{\xi^2} \frac{d^n}{d\xi^n} (e^{-\xi^2}), \quad (1.2.158)$$

which is the **Rodrigue's formula** for Hermite polynomials, but we do not know that yet and will need to show that it satisfies the equation. We proceed as follows:

$$\frac{\partial S}{\partial \xi} = 2se^{-s^2+2s\xi} = \sum \frac{2s^{n+1}}{n!} H_n, \quad (1.2.159)$$

but it is also equal to, based on Equation 1.2.157

$$\sum \frac{s^n}{n!} H'_n. \quad (1.2.160)$$

We do the same wrt s :

$$\frac{\partial S}{\partial s} = (-2s + 2\xi)e^{-s^2+2s\xi} = \sum \frac{(-2s + 2\xi)s^n}{n!} H_n = \sum \frac{s^{n-1}}{(n-1)!} H_n, \quad (1.2.161)$$

where the last equality is from Equation 1.2.157. Equating equal powers of s in the sums of $\partial S/\partial \xi$ and $\partial S/\partial s$ respectively, we have

$$H'_n = 2nH_{n-1} \quad (1.2.162a)$$

$$H_{n+1} = 2\xi H_n - 2nH_{n-1}. \quad (1.2.162b)$$

From Equations 1.2.162a and 1.2.162b we can construct:

$$H'_n = 2\xi H_n - H_{n+1} \quad (1.2.163a)$$

$$H''_n = 2H_n + 2\xi H'_n - H'_{n+1} \quad (1.2.163b)$$

$$H''_n = 2H_n + 2\xi H'_n - 2(n+1)H_n \quad (1.2.163c)$$

$$H''_n - 2\xi H'_n + 2nH_n = 0, \quad (1.2.163d)$$

the Hermite equation exactly. This shows that our definition in the forms of the generating function Equation 1.2.157 or the Rodrigue's formula Equation 1.2.158 do indeed correspond to Hermite polynomials.

We now pause and give a summary of the foregoing:

Definition 1.2.3.2 (Hermite polynomial). The Hermite polynomial of degree n is defined either by the generating function

$$\begin{aligned} S(\xi, s) &= S(\xi, s) = e^{\xi^2 - (s-\xi)^2} = e^{-s^2+2s\xi} \\ &\equiv \sum_{n=0}^{\infty} H_n \xi \frac{s^n}{n!}, \end{aligned} \quad (1.2.164)$$

which is not very useful, so the equivalent definition that is a result of a Taylor series expansion on the generating function is

$$H_n(\xi) = \frac{(-1)^n}{n!} e^{\xi^2} \frac{d^n}{d\xi^n} (e^{-\xi^2}). \quad (1.2.165)$$

Property 1.2.3.6 (Recurrence relation of Hermite polynomials).

$$H'_n = 2nH_{n-1} \quad (1.2.166a)$$

$$H_{n+1} = 2\xi H_n - 2nH_{n-1}. \quad (1.2.166b)$$

Now, on to the normalisation constant. We have

$$1 = \int_{-\infty}^{+\infty} |\psi|^2 dx = x_c |A_n|^2 \int_{-\infty}^{+\infty} H_n^2(\xi) e^{-\xi^2} d\xi. \quad (1.2.167)$$

The generating function now comes in handy as we can write the integral containing two generating functions:

$$\int_{-\infty}^{+\infty} S(\xi, s) S(\xi, t) e^{-\xi^2} d\xi = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{s^n t^m}{n! m!} \int_{-\infty}^{+\infty} H_n(\xi) H_m(\xi) e^{-\xi^2} d\xi, \quad (1.2.168)$$

where we can evaluate the LHS:

$$\int_{-\infty}^{+\infty} S(\xi, s) S(\xi, t) e^{-\xi^2} d\xi \quad (1.2.169a)$$

$$= \int_{-\infty}^{+\infty} e^{-s^2 + 2s\xi} e^{-t^2 + 2t\xi} e^{-\xi^2} d\xi \quad (1.2.169b)$$

$$= e^{-s^2 - t^2} \int_{-\infty}^{+\infty} e^{-\xi^2 + 2\xi(s+t)} d\xi \quad (1.2.169c)$$

$$= e^{-s^2 - t^2 + (s+t)^2} \int_{-\infty}^{+\infty} e^{-[\xi - (s+t)]^2} d\xi \quad (1.2.169d)$$

$$= e^{2st} \pi^{\frac{1}{2}} \quad (1.2.169e)$$

$$= \pi^{\frac{1}{2}} \sum_{n=0}^{\infty} \frac{(2st)^n}{n!}. \quad (1.2.169f)$$

This implies

$$\sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{s^n t^m}{n! m!} \int_{-\infty}^{+\infty} H_n(\xi) H_m(\xi) e^{-\xi^2} d\xi = \pi^{\frac{1}{2}} \sum_{k=0}^{\infty} \frac{(2st)^k}{k!}. \quad (1.2.169g)$$

Equating equal powers of s and t we get

$$\int_{-\infty}^{+\infty} H_n^2(\xi) e^{-\xi^2} d\xi = 2^n n! \sqrt{\pi} \quad (1.2.170a)$$

$$\int_{-\infty}^{+\infty} H_n(\xi) H_m(\xi) e^{-\xi^2} d\xi = 0. \quad (1.2.170b)$$

Beautiful. Now Equation 1.2.170a tells us we should choose, within an arbitrary complex

phase factor, the normalisation constant to be

$$A_n = \left(\frac{1}{x_c \pi^{\frac{1}{2}} 2^n n!} \right)^{\frac{1}{2}} = \left(\frac{m\omega}{\pi \hbar} \right)^{\frac{1}{4}} \frac{1}{\sqrt{2^n n!}}. \quad (1.2.171)$$

Finally, reverting back to dimensionful $\psi(x)$, we have the complete stationary state wave-function:

$$\psi(x) = \left(\frac{m\omega}{\pi \hbar} \right)^{\frac{1}{4}} \frac{1}{\sqrt{2^n n!}} H_n \left(\sqrt{\frac{m\omega}{\hbar}} x \right) e^{-m\omega x^2 / 2\hbar}. \quad (1.2.172)$$

That wasn't too bad.

1.2.3.3 The Morse potential

The morse potential is a more exact potential than the Hookeian potential, and it includes explicitly the effect of bond breaking: it approaches an asymptotic value, also known as the dissociation energy, D_e . As a result it admits scattering states.

Definition 1.2.3.3 (The Morse potential). The potential is given by

$$V_M(r) = D_e \left[1 - e^{-\beta(r-r_e)} \right]^2, \quad (1.2.173)$$

where r_e is the equilibrium bond length, and D_e is the dissociation energy, **not** bond strength as we will see soon.

Now if we expand the potential about $r = r_e$, to simply we rewrite

$$t \equiv \beta(r - r_e) \Rightarrow V_M = D_e(1 - e^{-t})^2. \quad (1.2.174)$$

So we have

$$\begin{aligned} V_M(r) &= D_e \left[1 - \left(1 - t + \frac{1}{2}t^2 - \frac{1}{6}t^3 + \dots \right) \right]^2 \\ &= D_e \left(t - \frac{1}{2}t^2 + \frac{1}{6}t^3 - \dots \right)^2 \\ &= D_e \left(t^2 - t^3 + \frac{7}{12}t^4 - \dots \right) \\ &= D_e \left[\beta^2(r - r_e)^2 - \beta^3(r - r_e)^3 + \dots \right]. \end{aligned} \quad (1.2.175)$$

So we can see from the quadratic term that, k_e , the force constant near equilibrium, *i.e.*, at small r about r_e where the quadratic term dominates, is given by

$$k_e = 2\beta^2 D_e. \quad (1.2.176)$$

Or it's to say that, β , the heretofore unspecified constant is

$$\beta = \sqrt{\frac{k_e}{2D_e}} \quad (1.2.177)$$

The Schrödinger equation is rather messy to solve:

$$\left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial r^2} + V(r) \right] \psi_n = E_n \psi_n \quad (1.2.178)$$

we can introduce some new variables to simplify the equation:

$$x \equiv ar, \quad x_e \equiv ar_e, \quad \lambda \equiv \frac{\sqrt{2mD_e}}{a\hbar}, \quad \epsilon_n \equiv \frac{2m}{a^2\hbar^2} E_n. \quad (1.2.179)$$

The Schrödinger equation now becomes

$$\left[-\frac{\partial^2}{\partial x^2} + V(x) \right] \psi_n(x) = \epsilon_n \psi_n(x), \quad (1.2.180)$$

where

$$V(x) = \lambda^2 \left[e^{-2(x-x_e)} - 2e^{-(x-x_e)} \right]. \quad (1.2.181)$$

todo: solve this or at least outline steps, priority:low.

An important result is very neat however: the eigenvalues, *i.e.*, allowed energies turn out to be

Property 1.2.3.7 (Eigenvalues of Morse potential).

$$E_n = \left(n + \frac{1}{2} \right) \hbar\omega - \left(n + \frac{1}{2} \right)^2 \hbar\omega x_e, \quad (1.2.182)$$

where $\omega = \sqrt{k_e/m}$ and

$$x_e = \frac{\hbar\beta^2}{2m\omega} = \frac{\hbar\omega}{4D_e} \quad (1.2.183)$$

is the **anharmonicity constant**.

Property 1.2.3.8 (Bond energy). The bond energy is the *difference* between the dissociation energy and the ground state energy:

$$E_B = D_e - E_0 = D_e - \frac{1}{2}\hbar\omega + \frac{1}{4}\hbar\omega x_e. \quad (1.2.184)$$

1.2.4 The free particle

This subsection deals with the case where

$$V(x) = 0. \quad (1.2.185)$$

Superficially this should be the simplest case of all, as the classical analogue is just constant velocity. However the case is not quite as straightforward as we shall see.

The Schrödinger equation reads

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} = E\psi, \quad (1.2.186)$$

or

$$\frac{d^2\psi}{dx^2} = -k^2\psi, \quad \text{where } k \equiv \frac{\sqrt{2mE}}{\hbar}. \quad (1.2.187)$$

We get the complex solution

$$\psi(x) = Ae^{ikx} + Be^{-ikx}. \quad (1.2.188)$$

Unlike the infinite square well, there are *no boundary conditions* - the free particle can carry any energy. The time-dependent solution is just

$$\Psi(x, t) = Ae^{ik(x-\hbar kt/2m)} + Be^{-ik(x+\hbar kt/2m)}, \quad (1.2.189)$$

which reminds us to write

$$\Psi(x, t) = \Psi(x \pm vt), \quad (1.2.190)$$

which describes a wave of fixed profile. We might as well write

$$\Psi_k(x, t) = Ae^{i(kx-\hbar k^2 t/2m)}, \quad \text{with } k \equiv \pm \frac{\sqrt{2mE}}{\hbar} \quad (1.2.191)$$

. The 'stationary states' have wavelengths

$$\lambda = \frac{2\pi}{|k|}, \quad (1.2.192)$$

and by de Broglie formula they carry momentum

$$p = \hbar k, \quad (1.2.193)$$

and velocity

$$v_q = \frac{\hbar|k|}{2m} = \sqrt{\frac{E}{2m}}. \quad (1.2.194)$$

However, classically, if we have energy E ,

$$v_c = \sqrt{\frac{2E}{m}} = 2v_q. \quad (1.2.195)$$

Not too good. But something way worse is here: these ‘stationary states’ are not normalisable:

$$\int_{-\infty}^{+\infty} \Psi_k^* \Psi_k dx = |A|^2 \int_{-\infty}^{+\infty} dx = |A|^2(\infty). \quad (1.2.196)$$

This means that there is no such thing as a free particle with a definite energy.

But wait, the complete time-dependent solution can be written, with a continuous variable coefficient $\phi(k)$:

$$\Psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k) e^{i(kx - \hbar k^2 t / 2m)} dk, \quad (1.2.197)$$

where the factor $1/\sqrt{2\pi}$ is factored out of $\phi(k)$ for convenience. This solution can be normalisable for appropriate choice of $\phi(k)$, but it necessarily carries a range of k ’s, and hence a range of energies and speeds. We term it a **wave packet**.

Returning to the generic quantum problem where we are given $\Psi(x, 0)$ and we need to find $\Psi(x, t)$, and we do so by determining the coefficient function $\phi(k)$:

$$\Psi(x, 0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k) e^{ikx} dk, \quad (1.2.198)$$

where the integrand is just the continuous version of the linear combination of stationary states, which is e^{ikx} in this problem.

But how to determine $\phi(k)$? We introduce

Theorem 1.2.4.1 (Plancherel’s theorem). Plancherel’s theorem states that

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} F(k) e^{ikx} dk \leftrightarrow F(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-ikx} dx. \quad (1.2.199)$$

$F(k)$ is called the **Fourier transform** of $f(x)$ and $f(x)$ is the **inverse Fourier transform** of $F(k)$, provided that the integrals exist.

Proof. Dirichlet’s theorem states that for any well behaved function $f(x)$ on the interval $[-a, +a]$,

$$f(x) = \sum_{n=0}^{\infty} [a_n \sin(n\pi x/a) + b_n \cos(n\pi x/a)]. \quad (1.2.200)$$

This can be written as a complex series

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{in\pi x/a}, \quad (1.2.201)$$

expanding we have

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{in\pi x/a} \quad (1.2.202a)$$

$$= c_0 + \sum_{n=1}^{\infty} (c_n + c_{-n}) \cos(n\pi x/a) + \sum_{n=1}^{\infty} i(c_n - c_{-n}) \sin(n\pi x/a) \quad (1.2.202b)$$

$$\equiv \sum_{n=0}^{\infty} [a_n \sin(n\pi x/a) + b_n \cos(n\pi x/a)]. \quad (1.2.202c)$$

So we have

$$c_n = \frac{1}{2}(b_n - ia_n) \quad (1.2.203a)$$

$$c_{-n} = \frac{1}{2}(b_n + ia_n). \quad (1.2.203b)$$

We can obtain the coefficients by

$$c_n = \frac{1}{2a} \int_{-a}^{+a} f(x) e^{-in\pi x/a} dx. \quad (1.2.204)$$

This is due to the orthogonality of $e^{-in\pi x/a}$:

$$\int_{-a}^{+a} e^{-in\pi x/a} e^{-im\pi x/a} dx = \frac{\delta_{nm}}{2a}. \quad (1.2.205)$$

We now introduce new variable

$$k \equiv \frac{n\pi}{a} \quad (1.2.206)$$

so

$$c_n(k) = \frac{1}{2a} \int_{-a}^{+a} f(x) e^{-ikx} dx. \quad (1.2.207)$$

We also introduce

$$F(k) \equiv \sqrt{\frac{2}{\pi}} a c_n(k), \quad (1.2.208)$$

such that

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{in\pi x/a} \quad (1.2.209a)$$

$$= \sum_{n=-\infty}^{\infty} c_n \Delta n e^{ikx} \quad (1.2.209b)$$

$$= \sum_{k=-\infty}^{\infty} \frac{F(k)}{a} \sqrt{\frac{\pi}{2}} \frac{a \Delta k}{\pi} e^{ikx} \quad (1.2.209c)$$

$$= \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} F(k) e^{ikx} \Delta k, \quad (1.2.209d)$$

and

$$F(k) = \frac{1}{\sqrt{2\pi}} \int_{-a}^{+a} f(x) e^{-ikx} dx. \quad (1.2.210)$$

In the limit that $a \rightarrow \infty$, $k \rightarrow 0$, we obtain

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} F(k) e^{ikx} dk \quad (1.2.211a)$$

$$F(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-ikx} dx. \quad (1.2.211b)$$

□

With Plancherel's theorem we can conclude that

$$\phi(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \Psi(x, 0) e^{-ikx} dx, \quad (1.2.212)$$

and we now see why a factor of $1/\sqrt{2\pi}$ was taken out.

Worked example 1.2.4.1. A free particle, which is initially localised in the range $-a < x < a$, is released at time $t = 0$:

$$\Psi(x, 0) = \begin{cases} A, & \text{if } -a < x < a, \\ 0, & \text{otherwise.} \end{cases} \quad (1.2.213)$$

Find $\Psi(x, t)$.

Solution:

First we normalised $\Psi(x, 0)$ to get

$$\Psi(x, 0) = \frac{1}{\sqrt{2a}}. \quad (1.2.214)$$

And then we calculate $\phi(k)$:

$$\phi(k) = \frac{1}{\sqrt{2a}} \frac{1}{\sqrt{2\pi}} \int_{-a}^{+a} e^{-ikx} dx \quad (1.2.215a)$$

$$= \frac{1}{2\sqrt{\pi a}} \left[\frac{e^{-ikx}}{-ik} \right]_{-a}^{+a} \quad (1.2.215b)$$

$$= \frac{1}{k\sqrt{\pi a}} \left(\frac{e^{ika} - e^{-ika}}{2i} \right) \quad (1.2.215c)$$

$$= \frac{1}{\sqrt{\pi a}} \frac{\sin(ka)}{k} \quad (1.2.215d)$$

$$= \sqrt{\frac{a}{\pi}} \text{sinc}(ka). \quad (1.2.215e)$$

Therefore we get the time-dependent solution

$$\begin{aligned} \Psi(x, t) &= \frac{1}{\pi\sqrt{2a}} \int_{-\infty}^{+\infty} \frac{\sin(ka)}{k} e^{i(kx - \hbar k^2 t / 2m)} dk \\ &= \frac{\sqrt{a/2}}{\pi} \int_{-\infty}^{+\infty} \text{sinc}(ka) e^{i(kx - \hbar k^2 t / 2m)} dk. \end{aligned} \quad (1.2.216)$$

This integral can't be generally solved except for a few cases. However we can observe some limiting behaviour:

At small a : The starting wavefunction is a localised spike, in fact, the Dirac delta function. Under small angle approximation, $\text{sinc}(\theta) = 1$

$$\phi(k) = \sqrt{\frac{a}{\pi}}. \quad (1.2.217)$$

It has no k dependence and is flat for all k . This means for an infinitesimally localised particle, the spread of momentum / velocity is infinite upon release.

At large a : The starting wavefunction is now a flat distribution and $\phi(k)$ is just a very sharp sinc function. This is exactly the reverse situation as above and illustrates again the uncertainty principle.

We now return to the problem of the quantum velocity being half the classical velocity. We introduce the following theorem

Theorem 1.2.4.2 (Group velocity). For a wave packet made of waves of many different frequencies, the group velocity, describing how fast the interference pattern, *i.e.*, *modulation* or *envelope* of the wave, moves, is given by

$$v_g = \frac{\partial \omega}{\partial k}, \quad (1.2.218)$$

where $\omega(k, \dots)$ is called the **dispersion relation**.

Proof. We need to determine the group velocity of a wave packet of the form

$$\Psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k) e^{i(kx - \omega t)} dk. \quad (1.2.219)$$

In this specific case,

$$\omega = \frac{\hbar k^2}{2m}, \quad (1.2.220)$$

but this proof is generally applicable.

Let's assume that $\phi(k)$ is narrowly peaked about k_0 , for it to have a narrow spread of speeds, *i.e.*, almost monochromatic, such that the shape of the packet does not change too rapidly. Since the integrand is negligible except around k_0 , we Taylor expand $\omega(k)$ about that point:

$$\omega(k) = \omega_0 + \omega'_0(k - k_0) + O(k^2). \quad (1.2.221)$$

We perform a change of variable

$$s \equiv k - k_0 \quad (1.2.222)$$

to give

$$\Psi(x, t) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k_0 + s) e^{i[(k_0 + s)x + (\omega_0 + \omega'_0 s)t]} ds \quad (1.2.223a)$$

$$= \frac{1}{\sqrt{2\pi}} e^{i(-\omega_0 t + k_0 \omega'_0 t)} \int_{-\infty}^{+\infty} \phi(k_0 + s) e^{i(k_0 + s)(x - \omega'_0 t)} ds. \quad (1.2.223b)$$

Nothing fancy there, just pure algebra. We can also see that when $t = 0$, we have

$$\Psi(x, 0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k_0 + s) e^{i(k_0 + s)x} ds. \quad (1.2.224)$$

The integrands in Equations 1.2.224 and 1.2.223b are the same except for the shift from x to $(x - \omega'_0 t)$, so

$$\Psi(x, t) \approx e^{i(-\omega_0 t + k_0 \omega'_0 t)} \Psi((x - \omega'_0 t), 0). \quad (1.2.225)$$

The complex factor in front is just a phase factor and will not affect $|\Psi|^2$, and from the

form of the function we see clearly that the wave packet moves at speed ω'_0 , which is to say

$$v_g = \frac{\partial \omega}{\partial k}. \quad (1.2.226)$$

This is in contrast with the phase velocity, of a specific component wave, given as

$$v_p = \frac{\omega}{k}. \quad (1.2.227)$$

□

1.2.5 The delta-function potential

1.2.5.1 Bound states and scattering states

Definition 1.2.5.1 (Bound and scattering states). In quantum mechanics, all the potentials we deal with go to 0 at $|\infty|$, the distinction between bound and scattering states is

$$\begin{cases} E < 0 \Rightarrow & \text{bound state,} \\ E > 0 \Rightarrow & \text{scattering state.} \end{cases} \quad (1.2.228)$$

Property 1.2.5.1. Bound states give rise to normalisable and quantised solutions whereas scattering states give unphysical and non-normalisable solutions.

1.2.5.2 The Dirac delta function

From this section onwards we explore potentials that give rise to both bound and scattering states.

The **Dirac delta function** is defined as

Definition 1.2.5.2 (Dirac delta function).

$$\delta(x) \equiv \begin{cases} 0, & \text{if } x \neq 0 \\ \infty, & \text{if } x = 0 \end{cases} \quad \text{with } \int_{-\infty}^{+\infty} \delta(x) \, dx = 1. \quad (1.2.229)$$

Property 1.2.5.2 (Extracting value under integral). We can write

$$f(x)\delta(x-a) = f(a)\delta(x-a), \quad (1.2.230)$$

so under integral,

$$\int_{-\infty}^{+\infty} f(x)\delta(x-a) dx = f(a) \int_{-\infty}^{+\infty} \delta(x-a) dx = f(a). \quad (1.2.231)$$

The following properties we prove will need this equality for two equivalent expressions D_1 and D_2 involving delta functions:

$$\int_{-\infty}^{+\infty} f(x)D_1(x) dx = \int_{-\infty}^{+\infty} f(x)D_2(x) dx. \quad (1.2.232)$$

Property 1.2.5.3 ('Wider' deltas).

$$\delta(cx) = \frac{1}{|c|}\delta(x). \quad (1.2.233)$$

Proof. We make the substitution

$$\alpha \equiv cx. \quad (1.2.234)$$

So we have

$$\int_{-\infty}^{+\infty} f(x)\delta(cx) dx = \frac{1}{c} \int_{-\infty}^{+\infty} f(\alpha/c)\delta(\alpha) d\alpha = \frac{1}{c}f(x) = \int_{-\infty}^{+\infty} f(x)\frac{1}{c}\delta(x) dx, \quad (1.2.235)$$

for $c > 0$, and for $c < 0$ we have

$$\int_{-\infty}^{+\infty} f(x)\delta(cx) dx = \frac{1}{c} \int_{+\infty}^{-\infty} f(\alpha/c)\delta(\alpha) d\alpha = -\frac{1}{c}f(x) = \int_{-\infty}^{+\infty} f(x)\frac{-1}{c}\delta(x) dx, \quad (1.2.236)$$

so we have by way of Equation 1.2.232 that

$$\delta(cx) = \frac{1}{|c|}\delta(x). \quad (1.2.237)$$

This is seemingly confusing but we have to realised that the 'scaled' delta gives a smaller but wider peak. \square

Property 1.2.5.4 (Derivative of Heaviside). The derivative of the Heaviside step function is the Dirac delta function.

Proof.

$$\int_{-\infty}^{+\infty} f(x) \frac{dH(x)}{dx} dx = f(0) \int_{-\infty}^{+\infty} \frac{dH(x)}{dx} dx = f(0)(1 - 0) = f(0), \quad (1.2.238)$$

which is the same behaviour as $\delta(x)$. \square

Property 1.2.5.5 (Fourier transform). This is not a rigorous mathematical result but only used for expediency in quantum mechanics:

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{ik(x'-x)} dk = \delta(x' - x). \quad (1.2.239)$$

Proof. Because we have that

$$\int_{-\infty}^{+\infty} \delta(x - x') f(x') dx' = f(x), \quad (1.2.240)$$

and from Theorem 1.2.4.1 we can also write

$$f(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-ikx} f(x) dx \quad (1.2.241a)$$

$$f(x') = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{ikx'} f(k) dx, \quad (1.2.241b)$$

substituting the top equation into the bottom one we have

$$f(x') = \int_{-\infty}^{+\infty} \left(\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{ik(x'-x)} dk \right) f(k) dx. \quad (1.2.242)$$

So we can identify

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{ik(x'-x)} dk = \delta(x' - x). \quad (1.2.243)$$

\square

1.2.5.3 Delta potential: bound states

Let's consider a potential of the form

$$V(x) = -\alpha \delta(x), \quad (1.2.244)$$

where $\alpha > 0$. The Schrödinger equation now reads

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} - \alpha\delta(x)\psi = E\psi. \quad (1.2.245)$$

We can see soon that this potential admits both bound and scattering states, *i.e.*, the energy can be both positive and negative in different parts of the solution.

We look for bound states first. In the region $x < 0$, $V(x) = 0$, so

$$\frac{d^2\psi}{dx^2} = \kappa^2\psi, \text{ where } \kappa \equiv \frac{\sqrt{-2mE}}{\hbar}, \quad (1.2.246)$$

where, as we are looking for bound states, $E < 0$. The solution is

$$\psi(x) = Ae^{-\kappa x} + Be^{\kappa x}, \quad (x < 0) \quad (1.2.247)$$

and we must have that $A = 0$ as the term will blow up as $x \rightarrow -\infty$ as we are concerned with the negative part of the potential. Now, for the positive part we similarly have

$$\psi(x) = Fe^{-\kappa x}, \quad (x > 0) \quad (1.2.248)$$

Now we just need to use the boundary conditions to ‘stitch’ the two solutions together. Continuity of wavefunction requires that

$$F = B. \quad (1.2.249)$$

So we get

$$\psi(x) = \begin{cases} Be^{\kappa x}, & (x \leq 0) \\ Be^{-\kappa x}, & (x \geq 0) \end{cases} \quad (1.2.250)$$

However, as the potential is infinite at $x = 0$ the continuity of first derivative does not need to be fulfilled. But the delta function has not come into the picture at all so far, and we need it to determine the discontinuity in the derivative. We do so by integrating the Schrödinger equation from $-\epsilon$ to $+\epsilon$ and take the limit as $\epsilon \rightarrow 0$:

$$-\frac{\hbar^2}{2m} \int_{-\epsilon}^{+\epsilon} \frac{d^2\psi}{dx^2} dx + \int_{-\epsilon}^{+\epsilon} V(x)\psi(x) dx = E \int_{-\epsilon}^{+\epsilon} \psi(x) dx \quad (1.2.251a)$$

$$\frac{2m}{\hbar^2} \lim_{\epsilon \rightarrow 0} \int_{-\epsilon}^{+\epsilon} V(x)\psi(x) dx = \lim_{\epsilon \rightarrow 0} \left(\frac{d\psi}{dx} \Big|_{+\epsilon} - \frac{d\psi}{dx} \Big|_{-\epsilon} \right) \equiv \Delta \left(\frac{d\psi}{dx} \right). \quad (1.2.251b)$$

This gives

$$\Delta \left(\frac{d\psi}{dx} \right) = -\frac{2m\alpha}{\hbar^2} \psi(0), \quad (1.2.252)$$

and for the present case we can readily evaluate $\Delta (d\psi/dx) = -2B\kappa$, so

$$\kappa = \frac{m\alpha}{\hbar^2}, \quad (1.2.253)$$

so the allowed energy turns out to be

$$E = -\frac{\hbar^2 \kappa^2}{2m} = -\frac{m\alpha^2}{2\hbar^2}, \quad (1.2.254)$$

the *only* allowed energy. We then normalise the wavefunction to get

$$B = \sqrt{\kappa} = \frac{\sqrt{m\alpha}}{\hbar}. \quad (1.2.255)$$

So we get the bound state wavefunction,

$$\psi(x) = \frac{\sqrt{m\alpha}}{\hbar} e^{-m\alpha|x|/\hbar^2}; \quad E = -\frac{m\alpha^2}{2\hbar^2}. \quad (1.2.256)$$

1.2.5.4 Delta potential: scattering states

For scattering states we solve, for $x < 0$,

$$\frac{d\psi}{dx} = -k^2\psi, \text{ where } k \equiv \frac{\sqrt{2mE}}{\hbar}. \quad (1.2.257)$$

The general solution is

$$\psi(x) = Ae^{ikx} + Be^{-ikx}, \quad (1.2.258)$$

similarly for $x > 0$,

$$\psi(x) = Fe^{ikx} + Ge^{-ikx}. \quad (1.2.259)$$

Continuity at $x = 0$ requires

$$F + G = A + B. \quad (1.2.260)$$

We also evaluate $\Delta (d\psi/dx)$ at $x = 0$:

$$\Delta \left(\frac{d\psi}{dx} \right) = ik(F - G - A + B). \quad (1.2.261)$$

We also have that $\psi(0) = (A + B)$, and according to Equation 1.2.252,

$$ik(F - G - A + B) = -\frac{2m\alpha}{\hbar^2}(A + B) \quad (1.2.262a)$$

$$F - G = A(1 + 2i\beta) - B(1 - 2i\beta), \text{ where } \beta \equiv \frac{m\alpha}{\hbar^2 k}. \quad (1.2.262b)$$

A and F is proportional to the amplitude of wave travelling to the right before and after the potential well (incident and transmitted), B and G for waves travelling to the left (reflected and right-incident). Now, restricting to a beam of particles coming from the left, we can have

$$G = 0. \quad (1.2.263)$$

Now the system of equations to solve is

$$A + B = F \quad (1.2.264a)$$

$$A(1 + 2i\beta) - B(1 - 2i\beta) = F. \quad (1.2.264b)$$

Solving, we have

$$B = \frac{i\beta}{1 - i\beta}A, \quad F = \frac{1}{1 - i\beta}A. \quad (1.2.264c)$$

The **reflection coefficient** is the relative probability that an incident particle *in a particle beam* will be reflected back:

$$R \equiv \frac{|\Psi_B|^2}{|\Psi_A|^2} = \frac{|A|^2}{|B|^2} = \frac{\beta^2}{1 + \beta^2}. \quad (1.2.265)$$

And the **transmission coefficient** is

$$T \equiv \frac{|F|^2}{|A|^2} = \frac{1}{1 + \beta^2}. \quad (1.2.266)$$

We note that R and T are functions of β^2 , where

$$\beta = \frac{m\alpha}{\hbar^2 k} = \frac{m\alpha^2}{2\hbar^2 E}. \quad (1.2.267)$$

However, as the free particle problem, the 'stationary states' are not normalisable, and must be realised in a wave packet with a range of energies. So the R and T should be treated as approximations near E .

Now notice how R and T are dependent on α^2 , so if the 'well' is now inverted to give a *barrier*, the particle beam is equally likely to tunnel through. However the bound state is now impossible due to Theorem 1.2.1.5.

1.3 FORMALISM

1.3.1 Hilbert space

Definition 1.3.1.1 (Hilbert space). A Hilbert space is a vector space consisting all **square-integrable functions** $f(x)$ on a specified interval such that

$$\int_a^b |f(x)|^2 dx < \infty. \quad (1.3.1)$$

In quantum mechanics we usually use the limits $[-\infty, +\infty]$.

Definition 1.3.1.2 (Inner product). The inner product of two functions $f(x)$ and $g(x)$ is defined

$$\langle f|g \rangle \equiv \int_a^b f^* g dx. \quad (1.3.2)$$

Theorem 1.3.1.1 (Schwarz inequality).

$$\left| \int_a^b f^* g dx \right| \leq \sqrt{\int_a^b |f|^2 dx \int_a^b |g|^2 dx}. \quad (1.3.3)$$

Proof. Consider the integral

$$I(a) = \int_a^b (af^* + g)(af + g^*) dx, \quad (1.3.4)$$

with a real. Expanding out we have

$$I(a) = \int_a^b a^2 |f|^2 + |g|^2 + a(f^* g + f g^*) dx \quad (1.3.5a)$$

$$= a^2 \langle f|f \rangle + \langle g|g \rangle + a(\langle f|g \rangle + \langle g|f \rangle) dx \quad (1.3.5b)$$

$$= a^2 \langle f|f \rangle + \langle g|g \rangle + a(\langle f|g \rangle + \langle f|g \rangle^*) dx \quad (1.3.5c)$$

Because $I(a)$ is necessarily non-negative, we calculate the discriminant wrt a ,

$$4|\langle f|g \rangle|^2 - 4\langle f|f \rangle \langle g|g \rangle \geq 0, \quad (1.3.6)$$

which gives the inequality we want.

We have used the fact that

$$\langle g|f \rangle = \langle f|g \rangle^*, \quad (1.3.7)$$

and that

$$\langle f|g\rangle^* \langle f|g\rangle = |\langle f|g\rangle|^2. \quad (1.3.8)$$

□

Definition 1.3.1.3 (Orthonormality). A function is said to be normalised if

$$\langle f|f\rangle = 1. \quad (1.3.9)$$

Two functions are orthonormal if

$$\langle f_m|f_n\rangle = \delta_{mn}. \quad (1.3.10)$$

Definition 1.3.1.4 (Completeness). A **set** of functions is complete if any other function in Hilbert space can be expressed as a linear combination of them:

$$f(x) = \sum_{n=1}^{\infty} c_n f_n(x), \quad (1.3.11)$$

if the f_n 's are orthonormal, the coefficients are given by

$$c_n = \langle f_n|f\rangle. \quad (1.3.12)$$

1.3.2 Observables

1.3.2.1 Hermitian operators

An observable quantity **must be real**, and so must be the average of many measurements. For an observable $Q(x, p)$ we can write

$$\langle Q \rangle = \int_{-\infty}^{+\infty} \Psi^* Q \Psi \, dx = \langle \Psi | Q \Psi \rangle. \quad (1.3.13)$$

The requirement that $\langle Q \rangle$ is real says that

$$\langle Q \rangle = \langle Q \rangle^* \quad (1.3.14a)$$

$$\langle \Psi | Q \Psi \rangle = \langle Q \Psi | \Psi \rangle, \quad (1.3.14b)$$

that is, Q must be a **Hermitian** operator. (Recall Lemma 1.2.3.1), we have proven the theorem

Theorem 1.3.2.1 (Observables correspond to hermitian operators). All physical observables correspond to hermitian operators.

Lemma 1.3.2.1 (Equivalent definition of Hermiticity). A Hermitian operator is equivalently defined by

$$\langle f|Qg\rangle = \langle Qf|g\rangle \quad (1.3.15)$$

for all $f(x)$ and all $g(x)$.

Proof. Let

$$h \equiv f + g, \quad (1.3.16)$$

noting that linear combinations of functions in Hilbert space give functions in Hilbert space, we have

$$\langle h|Qh\rangle = \langle Qh|h\rangle \quad (1.3.17a)$$

$$\langle (f + g)|(Qf + Qg)\rangle = \langle (Qf + Qg)|(f + g)\rangle \quad (1.3.17b)$$

$$\langle f|Qf\rangle + \langle f|Qg\rangle + \langle g|Qf\rangle + \langle g|Qg\rangle = \langle Qf|f\rangle + \langle Qf|g\rangle + \langle Qg|f\rangle + \langle Qg|g\rangle \quad (1.3.17c)$$

$$\langle f|Qg\rangle + \langle g|Qf\rangle = \langle Qf|g\rangle + \langle Qg|f\rangle. \quad (1.3.17d)$$

Now setting

$$h \equiv f + ig, \quad (1.3.18)$$

we can similarly write

$$\langle h|Qh\rangle = \langle Qh|h\rangle \quad (1.3.19a)$$

$$\langle (f + ig)|(Qf + iQg)\rangle = \langle (Qf + iQg)|(f + ig)\rangle \quad (1.3.19b)$$

$$\langle f|Qf\rangle + i\langle f|Qg\rangle - i\langle g|Qf\rangle + (-i)i\langle g|Qg\rangle = \langle Qf|f\rangle + i\langle Qf|g\rangle - i\langle Qg|f\rangle + (-i)i\langle Qg|g\rangle \quad (1.3.19c)$$

$$\langle f|Qg\rangle - \langle g|Qf\rangle = \langle Qf|g\rangle - \langle Qg|f\rangle. \quad (1.3.19d)$$

By comparing Equations 1.3.17a and 1.3.19a we immediately have

$$\langle f|Qg\rangle = \langle Qf|g\rangle, \quad (1.3.20)$$

for any $f(x)$ and $g(x)$ in the Hilbert space. \square

Theorem 1.3.2.2 (Commutator of two hermitian operators). The commutator of two hermitian operators is anti-hermitian, and same for two anti-hermitian operators.

Proof. Let P and Q be two hermitian operators, we can write

$$[P, Q]^\dagger = Q^\dagger P^\dagger - P^\dagger Q^\dagger = QP - PQ = [Q, P] = -[P, Q]. \quad (1.3.21)$$

For two anti-hermitian operators,

$$[P, Q]^\dagger = Q^\dagger P^\dagger - P^\dagger Q^\dagger = (-Q)(-P) - (-P)(-Q) = [Q, P] = -[P, Q]. \quad (1.3.22)$$

□

1.3.2.2 Determinate states

Definition 1.3.2.1 (Determinate states). A determinate state returns the same measurement q for the observable Q , in other words, the standard deviation of Q is zero.

Lemma 1.3.2.2 (Determinate states). Determinate states are eigenfunctions of Q .

Proof.

$$\sigma^2 = \langle (Q - \langle Q \rangle)^2 \rangle = \langle \Psi | (Q - q)^2 | \Psi \rangle = \langle (Q - q) \Psi | (Q - q) \Psi \rangle = 0. \quad (1.3.23)$$

This implies that

$$Q\Psi = q\Psi, \quad (1.3.24)$$

an **eigenvalue equation** for operator Q . □

Worked example 1.3.2.1. Consider the operator

$$Q \equiv i \frac{d}{d\phi}. \quad (1.3.25)$$

Is Q Hermitian? Find its eigenfunctions and eigenvalues.

Solution:

We are working with $f(\phi)$ on the finite interval $0 \leq \phi \leq 2\pi$, and for physical states we must have that

$$f(\phi + 2\pi) = f(\phi). \quad (1.3.26)$$

Using integration by parts

$$\langle f|Qg\rangle = \int_0^{2\pi} f^* \left(i \frac{dg}{d\phi} \right) d\phi \quad (1.3.27a)$$

$$= i f^* g \Big|_0^{2\pi} - \int_0^{2\pi} i \left(\frac{df^*}{d\phi} \right) g d\phi \quad (1.3.27b)$$

$$= \langle Qf|g\rangle, \quad (1.3.27c)$$

where the boundary terms only appeared because we imposed Equation 1.3.26. The eigenvalue equation is given by

$$i \frac{d}{d\phi} f(\phi) = q f(\phi). \quad (1.3.28)$$

The general solution is

$$f(\phi) = A e^{-iq\phi}. \quad (1.3.29)$$

Again, Equation 1.3.26 restricts us to possible values of q :

$$e^{-iq2\pi} = 1 \Rightarrow q = 0, \pm 1, \pm 2, \dots \quad (1.3.30)$$

It has a discrete spectrum of all integers and it is nondegenerate.

1.3.3 Eigenfunctions of a hermitian operator

1.3.3.1 Discrete spectra

Theorem 1.3.3.1 (Eigenvalues are real). Normalisable eigenfunctions of a hermitian operator have real eigenvalues.

Proof. We have

$$Qf = qf, \quad (1.3.31)$$

and

$$\langle f|Qf\rangle = \langle Qf|f\rangle, \quad (1.3.32)$$

so

$$q\langle f|f\rangle = q^*\langle f|f\rangle. \quad (1.3.33)$$

As

$$\langle f|f\rangle \neq 0 \quad (1.3.34)$$

by definition, we have

$$q = q^*. \quad (1.3.35)$$

□

Theorem 1.3.3.2 (Orthogonality). Eigenfunctions belonging to distinct eigenvalues are orthogonal.

Proof. We consider two eigenfunctions of Q :

$$Qf = qf \text{ and } Qg = q'g, \quad (1.3.36)$$

with Q hermitian, so

$$\langle f|Qg \rangle = \langle Qf|g \rangle, \quad (1.3.37)$$

and

$$q' \langle f|g \rangle = q^* \langle f|g \rangle. \quad (1.3.38)$$

Because q is real,

$$\langle f|g \rangle = 0. \quad (1.3.39)$$

□

For degenerate states, we can use the **Gram-Schmidt orthogonalisation procedure** to construct orthogonal eigenfunctions.

A further axiom is given that the eigenfunctions of an observable operator are complete, *i.e.*, any functions in Hilbert space can be expressed as a linear combination of them.

1.3.3.2 Continuous spectra

If the spectrum of a hermitian operator is continuous, the eigenfunctions are not normalisable in the sense we have used so far. We aim to resolve this.

We now inspect the momentum operator for its eigenfunctions, $f_p(x)$, and eigenvalues, p :

$$\frac{\hbar}{i} \frac{d}{dx} f_p(x) = p f_p(x). \quad (1.3.40)$$

The general solution is

$$f_p(x) = A e^{ipx/\hbar}. \quad (1.3.41)$$

For any, even *complex* values of p , these eigenfunctions are not square-integrable, hence the momentum operator has no eigenfunctions in Hilbert space, in the ‘traditional’ sense.

However, if we were to restrict ourselves to real eigenvalues, we can then write, with the help of Property 1.2.5.5 in the last equality,

$$\int_{-\infty}^{+\infty} f_{p'}^*(x) f_p(x) dx = |A|^2 \int_{-\infty}^{+\infty} e^{i(p'-p)x/\hbar} dx = |A|^2 2\pi\hbar \delta(p-p'). \quad (1.3.42)$$

We should evidently pick $A = 1/\sqrt{2\pi\hbar}$, so that

$$f_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar}. \quad (1.3.43)$$

We have

Definition 1.3.3.1 (Dirac orthonormality). A function is Dirac-orthonormalised if

$$\langle f_{p'} | f_p \rangle = \delta(p-p'). \quad (1.3.44)$$

Property 1.3.3.1 (Completeness of Dirac orthonormalised functions). These eigenfunctions are complete: for any function $f(x)$

$$f(x) = \int_{-\infty}^{+\infty} c(p) f_p(x) dp = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} c(p) e^{ipx/\hbar} dp, \quad (1.3.45)$$

with the ‘coefficient’ function $c(p)$ evaluated as follows:

$$\langle f_{p'} | f \rangle = \int_{-\infty}^{+\infty} c(p) \langle f_{p'} | f_p \rangle dp = \int_{-\infty}^{+\infty} c(p) \delta(p-p') dp = c(p'). \quad (1.3.46)$$

Looked at another way these equations just describe Fourier transforms between $f(x)$ and $c(p)$.

Therefore we see that the eigenfunctions of momentum are sinusoidal, with wavelengths

$$\lambda = \frac{2\pi\hbar}{p} = \frac{h}{p}. \quad (1.3.47)$$

This is the de Broglie formula, but with an important point to note: as the eigenfunctions are non-normalisable, a particle cannot have a determinate momentum, hence all we could talk about is a normalisable wave packet with a narrow range of momenta.

We now look at the position operator.

Let $g_y(x)$ be its eigenfunction and y be the eigenvalue:

$$xg_y(x) = yg_y(x), \quad (1.3.48)$$

where y is a fixed number but x is a continuous variable. We can identify immediately that

$$g_y(x) = A\delta(x - y). \quad (1.3.49)$$

The eigenfunctions are Dirac orthonormal:

$$\int_{-\infty}^{+\infty} g_{y'}^*(x) g_y(x) dx = |A|^2 \int_{-\infty}^{+\infty} \delta(x - y') \delta(x - y) dx = |A|^2 \delta(y - y'). \quad (1.3.50)$$

Picking $A = 1$,

$$g_y(x) = \delta(x - y). \quad (1.3.51)$$

And

$$\langle g_{y'} | g_y \rangle = \delta(y - y'). \quad (1.3.52)$$

They are also complete:

$$f(x) = \int_{-\infty}^{+\infty} c(y) g_y(x) dy = \int_{-\infty}^{+\infty} c(y) \delta(x - y) dy, \quad (1.3.53)$$

and we can identify

$$c(y) = f(y). \quad (1.3.54)$$

1.3.4 Generalised statistical interpretation

Postulate 1.3.4.1 (Generalised statistical interpretation). If you measure an observable $Q(x, p)$ on a particle in the state $\Psi(x, t)$, you can only get one of the eigenvalues of the hermitian operator $\hat{Q}(x, -i\hbar d/dx)$.

If \hat{Q} has a discrete spectrum, the probability of getting eigenvalue q_n of the eigenfunction $f_n(x)$ is

$$|c_n|^2, \text{ where } c_n = \langle f_n | \Psi \rangle. \quad (1.3.55)$$

If \hat{Q} has a continuous spectrum, with real eigenvalues $q(z)$ and associated Dirac-orthonormalised eigenfunctions $f_z(x)$, the probability of getting a result in the range dz is

$$|c(z)|^2 dz, \text{ where } c(z) = \langle f_z | \Psi \rangle. \quad (1.3.56)$$

Corollary 1.3.4.1 (Position measurement). We try to recover the formula we've been using for position measurements from the generalised statistical interpretation. We have the eigenfunction of position operator

$$g_y(x) = \delta(x - y), \quad (1.3.57)$$

so

$$c(y) = \langle f_y | \Psi \rangle = \int_{-\infty}^{+\infty} \delta(x - y) \Psi(x, t) dx = \Psi(y, t), \quad (1.3.58)$$

so the probability of getting a result in the range of dy is $|\Psi(y, t)|^2 dy$, which recovers the formula nicely.

Corollary 1.3.4.2 (Momentum space wavefunction). The eigenfunctions of momentum operator is

$$f_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar}, \quad (1.3.59)$$

so the coefficients are

$$c(p) = \langle f_p | \Psi \rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{-ipx/\hbar} \Psi(x, t) dx \quad (1.3.60)$$

$c(p, t)$ here is important enough to get another name and symbol: the **momentum space wavefunction** $\Phi(p, t)$:

$$\Phi(p, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{-ipx/\hbar} \Psi(x, t) dx. \quad (1.3.61)$$

This is in fact just a Fourier transform of $\Psi(x, t)$. The probability that a measurement of momentum would yield a result in the range dp is (note that we are working in momentum space now):

$$|\Phi(x, t)|^2 dp. \quad (1.3.62)$$

Corollary 1.3.4.3 (Momentum space position and momentum operators). The momentum space position operator is

$$\hat{X} \equiv -\frac{\hbar}{i} \frac{\partial}{\partial p}, \quad (1.3.63)$$

and the momentum space position operator is

$$\hat{P} \equiv p. \quad (1.3.64)$$

Proof. For the position operator we first note that

$$-\frac{\hbar}{i} \frac{\partial}{\partial p} \Phi = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} x e^{-ipx/\hbar} \Psi dx, \quad (1.3.65)$$

so that

$$\int_{-\infty}^{+\infty} \Phi^* \left(-\frac{\hbar}{i} \frac{\partial}{\partial p} \right) \Phi dp \quad (1.3.66a)$$

$$= \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} e^{ipx'/\hbar} \Psi^*(x') dx' \right) \left(\int_{-\infty}^{+\infty} x e^{-ipx/\hbar} \Psi(x) dx \right) dp \quad (1.3.66b)$$

$$= \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{ip(x'-x)/\hbar} \Psi^*(x') x \Psi(x) dx' dx \right) dp \quad (1.3.66c)$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \delta(x' - x) \Psi^*(x') x \Psi(x) dx' dx \quad (1.3.66d)$$

$$= \int_{-\infty}^{+\infty} \Psi^*(x) x \Psi(x) dx \quad (1.3.66e)$$

$$= \langle x \rangle. \quad (1.3.66f)$$

The momentum operator can be proven similarly. \square

Theorem 1.3.4.1 (Transform from position to momentum space).

Generally,

$$\langle Q(x, p) \rangle = \begin{cases} \int_{-\infty}^{+\infty} \Psi^* \hat{Q} \left(x, \frac{\hbar}{i} \frac{\partial}{\partial x} \right) \Psi dx, & \text{in position space;} \\ \int_{-\infty}^{+\infty} \Phi^* \hat{Q} \left(-\frac{\hbar}{i} \frac{\partial}{\partial p}, p \right) \Phi dp, & \text{in momentum space.} \end{cases} \quad (1.3.67)$$

1.3.5 The uncertainty principle

1.3.5.1 Proof of the generalised uncertainty principle

For any observable A we have

$$\sigma_A^2 = \langle \Psi | (\hat{A} - \langle A \rangle)^2 | \Psi \rangle = \langle (\hat{A} - \langle A \rangle) \Psi | (\hat{A} - \langle A \rangle) \Psi \rangle \equiv \langle f | f \rangle, \quad (1.3.68)$$

where we have used the hermiticity of \hat{A} and consequently of $\hat{A} - \langle A \rangle$ in the second equality.

Likewise, for any other observable B for which we define $g \equiv (\hat{B} - \langle B \rangle)\Psi$,

$$\sigma_B^2 = \langle g|g \rangle. \quad (1.3.69)$$

Therefore, invoking the Schwarz inequality we can write

$$\sigma_A^2 \sigma_B^2 = \langle f|f \rangle \langle g|g \rangle \geq |\langle f|g \rangle|^2. \quad (1.3.70)$$

Now for any complex number z ,

$$|z|^2 = \text{Re}(z)^2 + \text{Im}(z)^2 \geq \text{Im}(z)^2 = \left[\frac{1}{2i}(z - z^*) \right]^2. \quad (1.3.71)$$

Therefore, letting $z = \langle f|g \rangle$,

$$\sigma_A^2 \sigma_B^2 \geq \left(\frac{1}{2i} [\langle f|g \rangle - \langle g|f \rangle] \right)^2. \quad (1.3.72)$$

Now it's time to evaluate

$$\langle f|g \rangle = \langle (\hat{A} - \langle A \rangle)\Psi | (\hat{B} - \langle B \rangle)\Psi \rangle \quad (1.3.73a)$$

$$= \langle \Psi | (\hat{A} - \langle A \rangle)(\hat{B} - \langle B \rangle) \Psi \rangle \quad (1.3.73b)$$

$$= \langle \Psi | (\hat{A}\hat{B} - \hat{A}\langle B \rangle - \hat{B}\langle A \rangle + \langle A \rangle\langle B \rangle) \Psi \rangle \quad (1.3.73c)$$

$$= \langle \hat{A}\hat{B} \rangle - \langle B \rangle\langle A \rangle - \langle A \rangle\langle B \rangle + \langle A \rangle\langle B \rangle \quad (1.3.73d)$$

$$= \langle \hat{A}\hat{B} \rangle - \langle A \rangle\langle B \rangle. \quad (1.3.73e)$$

Similarly,

$$\langle g|f \rangle = \langle \hat{B}\hat{A} \rangle - \langle A \rangle\langle B \rangle. \quad (1.3.74)$$

So,

$$\langle f|g \rangle - \langle g|f \rangle = \langle \hat{A}\hat{B} \rangle - \langle \hat{B}\hat{A} \rangle = \langle [\hat{A}, \hat{B}] \rangle. \quad (1.3.75)$$

So we reach the conclusion:

Theorem 1.3.5.1 (Generalised uncertainty principle).

$$\sigma_A^2 \sigma_B^2 \geq \left(\frac{1}{2i} \langle [\hat{A}, \hat{B}] \rangle \right)^2. \quad (1.3.76)$$

We now prove a couple commutator identities:

Property 1.3.5.1 (Associativity and distributivity). Commutators are associative:

$$[AB, C] = A[B, C] + [A, C]B, \quad (1.3.77a)$$

$$[A, BC] = B[A, C] + [A, B]C. \quad (1.3.77b)$$

and distributive with respect to addition:

$$[A, B + C] = [A, B] + [A, C]. \quad (1.3.78)$$

Proof. For the first one we can write the LHS as

$$(ABC - CAB), \quad (1.3.79)$$

and the RHS

$$A(BC - CB) + (AC - CA)B = ABC - ACB + ACB - CAB = ABC - CAB = L.H.S. \quad (1.3.80)$$

$[A, BC]$ follows by anticommutativity. For the second one,

$$[A, B + C] = A(B + C) - (B + C)A = [A, B] + [A, C]. \quad (1.3.81)$$

□

Property 1.3.5.2 (Commutator with momentum).

Generally for any function $f(x)$,

$$[f(x), p] = i\hbar \frac{df}{dx}. \quad (1.3.82)$$

Proof.

$$[f(x), p]g = -f i\hbar \frac{dg}{dx} + i\hbar \frac{df}{dx} g \quad (1.3.83a)$$

$$= i\hbar \frac{df}{dx} g. \quad (1.3.83b)$$

□

Worked example 1.3.5.1. Derive the uncertainty principle of position and energy.

Solution:

We have, from the generalised uncertainty principle that

$$\sigma_x^2 \sigma_H^2 = \left(\frac{1}{2i} \langle [\hat{X}, \hat{H}] \rangle \right)^2. \quad (1.3.84)$$

Letting the commutator act on a test function $f(x)$, we have

$$[X, H] = -\frac{\hbar^2}{2m} \left[x \frac{d^2 f}{dx^2} + x V(x) f - \frac{d^2(xf)}{dx^2} - V(x) x f \right] \quad (1.3.85a)$$

$$= -\frac{\hbar^2}{2m} \left(-2 \frac{df}{dx} \right) \quad (1.3.85b)$$

$$= \frac{\hbar^2}{m} \frac{df}{dx}. \quad (1.3.85c)$$

So the uncertainty principle is

$$\sigma_x^2 \sigma_H^2 \geq \frac{\hbar^2}{4m^2} \langle p \rangle^2 \Rightarrow \sigma_x \sigma_H \geq \frac{\hbar}{2m} |\langle p \rangle|. \quad (1.3.86)$$

This relation is not useful for stationary states as $\langle p \rangle = 0$, but can prove useful for linear combinations of stationary states, when the average is non-zero.

Theorem 1.3.5.2 (Eigenfunctions of noncommuting operators). Two noncommuting operators cannot have a complete set of common eigenfunctions.

Proof. Let P and Q have a complete, i.e., any function can be written as a linear combination of them, set of common eigenfunctions, noting that they can have different eigenvalues, we'll have, for a function g in Hilbert space,

$$g = \sum c_n f_n. \quad (1.3.87)$$

But because we know that

$$P f_n = p_n f_n, \quad Q f_n = q_n f_n, \quad (1.3.88)$$

so

$$PQg = QPg = \sum c_n p_n q_n f_n. \quad (1.3.89)$$

So

$$[P, Q] = 0. \quad (1.3.90)$$

□

A related theorem can be stated as well

Theorem 1.3.5.3 (Eigenfunctions of commuting operators). Two commuting operators share a simultaneous set of eigenstates. todo-supo: check what this means

Proof. Let operators A and B commute, and for A we have

$$A|f_n\rangle = a_n|f_n\rangle, \quad (1.3.91)$$

right-multiplying by B we have

$$BA|f_n\rangle = a_nB|f_n\rangle \equiv a_n|\beta_n\rangle. \quad (1.3.92)$$

Now the commutator tells us that

$$BA|f_n\rangle = AB|f_n\rangle = A|\beta_n\rangle. \quad (1.3.93)$$

Comparing, we have

$$A|\beta_n\rangle = a_n|\beta_n\rangle. \quad (1.3.94)$$

This can mean one of two things:

(a) $|\beta_n\rangle$ is an eigenfunction of A with eigenvalue a_n , which implies

$$|\beta_n\rangle = B|f_n\rangle = b_n|f_n\rangle. \quad (1.3.95)$$

This means the two operators have a common set of eigenfunctions.

(b) todo/todo-supo: check the degenerate case, priority: medium [16], [12] □

1.3.5.2 The minimum-uncertainty wave packet

The harmonic oscillator and the Gaussian wave packet for free particle were two previous cases that hit the minimum uncertainty. We now ask when, in general, the equality holds for the uncertainty relation.

The Schwarz inequality, which was used in proving the uncertainty principle, becomes an equality if and only if one function is a multiple of the other, *i.e.*,

$$g(x) = cf(x), \quad (1.3.96)$$

where c is some complex number. Meanwhile, if Equation 1.3.71 were to become an equality, we must also require that

$$\text{Re}(\langle f|g\rangle) = 0. \quad (1.3.97)$$

Combining the two conditions we must have

$$\text{Re}(c\langle f|f\rangle) = 0. \quad (1.3.98)$$

However, $\langle f|f\rangle$ is necessarily real, which means c must be purely imaginary, which is to say

$$c \equiv ia, \quad (1.3.99)$$

where a is real. The condition for minimum uncertainty is then

$$g(x) = ia f(x). \quad (1.3.100)$$

Now recalling the definitions of g and f from Equation 1.3.68, this time for momentum-position uncertainty principle:

$$f = \hat{P} - \langle p \rangle = \frac{\hbar}{i} \frac{d}{dx} - \langle p \rangle \quad (1.3.101a)$$

$$g = x - \langle x \rangle. \quad (1.3.101b)$$

Therefore we have the resulting differential equation:

$$\left(\frac{\hbar}{i} \frac{d}{dx} - \langle p \rangle \right) \Psi = ia(x - \langle x \rangle) \Psi \quad (1.3.102a)$$

$$\frac{d\Psi}{dx} = \frac{i}{\hbar} [ia(x - \langle x \rangle) + \langle p \rangle] \Psi \quad (1.3.102b)$$

$$\frac{d\Psi}{dx} \equiv f(x) \Psi. \quad (1.3.102c)$$

The general solution is then

$$\Psi(x) = A e^{-a(x - \langle x \rangle)^2 / 2\hbar} e^{i\langle p \rangle x / \hbar}. \quad (1.3.103)$$

This is an instantaneous waveform as A , a , $\langle x \rangle$, $\langle p \rangle$ can all have time dependence, and the only thing this proves is that at *some* instant in time, the wavefunction can reach minimum uncertainty, and that the waveform at this instant is *Gaussian*.

1.3.5.3 The energy-time uncertainty principle

Theorem 1.3.5.4 (E-t uncertainty principle). The energy-time uncertainty principle states that

$$\Delta t \Delta E \geq \frac{\hbar}{2}. \quad (1.3.104)$$

This may look familiar, and indeed this is an immediate consequence of the x - p uncertainty principle, if we're working in relativistic quantum mechanics. But we're not, and it results from very different principles at work. This requires a proof.

Proof. For some observable $Q(x, p, t)$ (notice the new time dependence), we can compute the time derivative:

$$\frac{d}{dt}\langle Q \rangle = \frac{d}{dt}\langle \Psi | \hat{Q} | \Psi \rangle = \left\langle \frac{\partial \Psi}{\partial t} | \hat{Q} | \Psi \right\rangle + \left\langle \Psi | \frac{\partial \hat{Q}}{\partial t} | \Psi \right\rangle + \left\langle \Psi | \hat{Q} | \frac{\partial \Psi}{\partial t} \right\rangle. \quad (1.3.105)$$

Because the Schrödinger equation says

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H} \Psi, \quad (1.3.106)$$

we can simplify the time derivative

$$\frac{d}{dt}\langle Q \rangle = -\frac{1}{i\hbar}\langle \hat{H} \Psi | \hat{Q} | \Psi \rangle + \frac{1}{i\hbar}\langle \Psi | \hat{Q} | \hat{H} \Psi \rangle + \left\langle \frac{\partial \hat{Q}}{\partial t} \right\rangle. \quad (1.3.107)$$

The Hamiltonian is hermitian, so

$$\frac{d}{dt}\langle Q \rangle = \frac{i}{\hbar}\langle [\hat{H}, \hat{Q}] \rangle + \left\langle \frac{\partial \hat{Q}}{\partial t} \right\rangle. \quad (1.3.108)$$

We will use this in Corollary 1.3.5.1 and Theorem 1.3.5.5 to prove a few very useful theorems. Now, invoking the generalised uncertainty principle, setting $A = H$ and $B = Q$, and *assuming that Q does not depend on t* , we have

$$\sigma_H^2 \sigma_Q^2 \geq \left(\frac{1}{2i} \langle [\hat{H}, \hat{Q}] \rangle \right)^2 = \left(\frac{1}{2i} \hbar \frac{d\langle Q \rangle}{dt} \right)^2, \quad (1.3.109)$$

which is to say

$$\sigma_H \sigma_Q \geq \frac{\hbar}{2} \left| \frac{d\langle Q \rangle}{dt} \right|. \quad (1.3.110)$$

We now make the following definitions:

$$\Delta E \equiv \sigma_H \quad (1.3.111a)$$

$$\Delta t \equiv \frac{\sigma_Q}{|d\langle Q \rangle/dt|}. \quad (1.3.111b)$$

Δt here does *not* mean the uncertainty in time as it is not a *dynamical variable*, however we can see that

$$\sigma_Q = \left| \frac{d\langle Q \rangle}{dt} \right| \Delta t, \quad (1.3.112)$$

therefore Δt represents the time it takes the expectation value of Q to change by one standard deviation. We arrive at the E - t uncertainty principle:

$$\Delta E \Delta t \geq \frac{\hbar}{2}. \quad (1.3.113)$$

□

Corollary 1.3.5.1 (Time derivative of expectation/Heisenberg equation). From Equation 1.3.108 we have a useful corollary:

$$\frac{d}{dt} \langle Q \rangle = \frac{i}{\hbar} \langle [\hat{H}, \hat{Q}] \rangle + \left\langle \frac{\partial \hat{Q}}{\partial t} \right\rangle. \quad (1.3.114)$$

We apply it to several cases and appreciate its utility.

(a) With $Q = 1$, we know that all operators commute with a constant, and Q has no dependence on time, so

$$\frac{d\langle Q \rangle}{dt} = 0. \quad (1.3.115)$$

Looked at another way,

$$\langle Q \rangle = \langle \Psi | \Psi \rangle, \quad (1.3.116)$$

and the fact that time derivative is zero is simply a restatement of preservation of normalisation theorem (Theorem 1.1.1.1).

(b) With $Q = H$, since any operator commutes with itself, *i.e.*, the commutator is 0 we have

$$\frac{d\langle H \rangle}{dt} = \left\langle \frac{\partial H}{\partial t} \right\rangle. \quad (1.3.117)$$

If the Hamiltonian has no explicit energy dependence, *i.e.*, the potential is time-independent, then this is a restatement of conservation of energy.

(c) With $Q = x$, we already know the commutator from Equation 1.3.85c:

$$[\hat{H}, \hat{X}] = -[\hat{H}, \hat{X}] = -\frac{\hbar^2}{m} \frac{\partial}{\partial x} = -\frac{i\hbar}{m} \hat{P}. \quad (1.3.118)$$

As x is independent of time,

$$\frac{d\langle x \rangle}{dt} = \frac{i}{\hbar} \langle [\hat{H}, \hat{X}] \rangle = \frac{\langle p \rangle}{m}. \quad (1.3.119)$$

This is a restatement of the first part of Ehrenfest theorem (Equation 1.1.14a).

(d) With $Q = P$, we need to work out the commutator first:

$$[\hat{H}, \hat{P}]g = \hat{H} \left[\frac{\hbar}{i} \frac{\partial g}{\partial x} \right] - \hat{P} \left[-\frac{\hbar^2}{2m} \frac{\partial^2 g}{\partial x^2} + Vg \right] \quad (1.3.120a)$$

$$= -\frac{\hbar^2}{2mi} \frac{\partial^3 g}{\partial x^3} + \frac{\hbar}{i} V \frac{\partial g}{\partial x} + \frac{\hbar^2}{2mi} \frac{\partial^3 g}{\partial x^3} - \frac{\hbar}{i} \left(g \frac{\partial V}{\partial x} + V \frac{\partial g}{\partial x} \right) \quad (1.3.120b)$$

$$= i\hbar \frac{\partial V}{\partial x} g. \quad (1.3.120c)$$

Therefore,

$$\frac{d\langle p \rangle}{dt} = - \left\langle \frac{\partial V}{\partial x} \right\rangle. \quad (1.3.121)$$

This is a restatement of Ehrenfest theorem.

Theorem 1.3.5.5 (Virial theorem). The quantum mechanical Virial theorem states that

$$2\langle T \rangle = \left\langle x \frac{dV}{dx} \right\rangle, \quad (1.3.122)$$

where T is the kinetic energy.

Proof. We use Corollary 1.3.5.1 to prove this theorem. We can write, for $Q \equiv XP$, the commutator

$$[\hat{H}, \hat{X}\hat{P}]f = \left(-\frac{\hbar^2}{2m} \frac{\partial}{\partial x} + V(x) \right) \left(x \frac{\hbar}{i} \frac{\partial}{\partial x} \right) f - \left(x \frac{\hbar}{i} \frac{\partial}{\partial x} \right) \left(-\frac{\hbar^2}{2m} \frac{\partial}{\partial x} + V(x) \right) f \quad (1.3.123a)$$

$$= -\frac{\hbar^3}{im} \frac{\partial^2 f}{\partial x^2} - \frac{\hbar}{i} x \frac{\partial V}{\partial x} f. \quad (1.3.123b)$$

Corollary 1.3.5.1 gives

$$\frac{d}{dt} \langle xp \rangle = \frac{i}{\hbar} \langle [H, XP] \rangle + \left\langle \frac{\partial(xp)}{\partial t} \right\rangle. \quad (1.3.124)$$

The last term is zero because the operator $\hat{X}\hat{P}$ has no time dependence (see Footnote 7 of [20]). The LHS can be re-written

$$\langle xp \rangle = \langle \Psi | XP | \Psi \rangle = \langle \Psi | X | \Psi \rangle \langle \Psi | P | \Psi \rangle = \langle x \rangle \langle p \rangle, \quad (1.3.125)$$

therefore, in stationary states,

$$\frac{d}{dt} \langle xp \rangle = \langle x \rangle' \langle p \rangle + \langle x \rangle \langle p \rangle' = 0. \quad (1.3.126)$$

Therefore

$$0 = \frac{i}{\hbar} \langle [H, XP] \rangle \quad (1.3.127a)$$

$$= \left\langle -\frac{\hbar^2}{m} \frac{\partial^2}{\partial x^2} - x \frac{\partial V}{\partial x} \right\rangle \quad (1.3.127b)$$

$$= 2\langle T \rangle - \left\langle x \frac{dV}{dx} \right\rangle. \quad (1.3.127c)$$

Remember that we have assumed that the potential is time-independent. This gives the required theorem. \square

Corollary 1.3.5.2 (Virial theorem for harmonic oscillator). For harmonic oscillators we have

$$V(x) = \frac{1}{2} m \omega^2 x^2 \quad (1.3.128)$$

So the Virial theorem gives us

$$2\langle T \rangle = \langle m \omega^2 x^2 \rangle = 2\langle V \rangle \Rightarrow \langle T \rangle = \langle V \rangle. \quad (1.3.129)$$

1.3.6 Dirac Notation

come back and do the exercises also, priority: medium

Worked example 1.3.6.1. In this problem we deal with **sequential measurements**. Suppose an observable operator A has two normalised eigenstates ψ_1 and ψ_2 , and the operator B has ϕ_1 and ϕ_2 , and they are related by (remember that both must have a *complete* set of eigenstates, and as such they can *always* be expressed in terms of each other.)

$$\psi_1 = (3\phi_1 + 4\phi_2)/5, \quad \psi_2 = (4\phi_1 - 3\phi_2)/5. \quad (1.3.130)$$

(a) Observable A is measured and the value a_1 is obtained, what is the state of the system immediately after the measurement?

Solution: The system must now be in state ψ_1 as the wavefunction has collapsed.

(b) If B is now measured, what are the possible results and their probabilities?

Solution: As the system is now in the state $\psi_1 = (3\phi_1 + 4\phi_2)/5$, the probability of getting either b_1 or b_2 is $9/25$ or $16/25$.

(c) Right after the measurement of B , A is measured again. What's the probability of getting a_1 again?

Solution: We compute the conditional probabilities

$$P(a_1) = P(a_1|b_1)P(b_1) + P(a_1|b_2)P(b_2). \quad (1.3.131)$$

To get the conditional probabilities we need to invert the relations:

$$\phi_1 = (3\psi_1 + 4\psi_2)/5, \quad \phi_2 = (4\psi_1 - 3\psi_2)/5, \quad (1.3.132)$$

so

$$P(a_1) = \frac{9}{25} \cdot \frac{9}{25} + \frac{16}{25} \cdot \frac{16}{25} \approx 0.54. \quad (1.3.133)$$

1.4 QUANTUM MECHANICS IN THREE DIMENSIONS

1.4.1 Schrödinger equation in 3D with non-Coulombic potentials

1.4.1.1 Preparatory work

We introduce the 3D Schrödinger equation as follows:

$$i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}, t) = -\frac{\hbar^2}{2m} \nabla^2 \Psi(\mathbf{r}, t) + V\Psi(\mathbf{r}, t), \quad (1.4.1)$$

where

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \quad (1.4.2)$$

Property 1.4.1.1 (Canonical commutation relations in 3D). We have

$$[r_i, p_j] = -[p_i, r_j] = i\hbar \delta_{ij}, \quad [r_i, r_j] = [p_i, p_j] = 0. \quad (1.4.3)$$

The proof is trivial.

Corollary 1.4.1.1 (Heisenberg's uncertainty in 3D). It follows from the canonical commutator relations in 3D that

$$\sigma_{x_i} \sigma_{p_j} \geq \frac{\hbar}{2} \delta_{ij}. \quad (1.4.4)$$

Theorem 1.4.1.1 (Ehrenfest's theorem in 3D).

$$\frac{d}{dt} \langle \mathbf{r} \rangle = \frac{1}{m} \langle \mathbf{p} \rangle \quad (1.4.5a)$$

$$\frac{d}{dt} \langle \mathbf{p} \rangle = \langle -\nabla V \rangle. \quad (1.4.5b)$$

This is true because in deriving Equation 1.3.108, no calculation depended on the dimensions, so we can do the same calculation in all three cardinal directions separately and get the 3D theorem.

Property 1.4.1.2 (Laplacian in spherical coordinates). The Laplacian ∇^2 in spherical coordinates is

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \left(\frac{\partial^2}{\partial \phi^2} \right). \quad (1.4.6)$$

Proof. It can be derived either from a tedious change-of-basis partial differentiation exercise or from generalised curvilinear coordinates, which we will use:

The Laplacian in curvilinear coordinates is

$$\frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial q_1} \left(\frac{h_2 h_3}{h_1} \frac{\partial}{\partial q_1} \right) + \frac{\partial}{\partial q_2} \left(\frac{h_3 h_1}{h_2} \frac{\partial}{\partial q_2} \right) + \frac{\partial}{\partial q_3} \left(\frac{h_1 h_2}{h_3} \frac{\partial}{\partial q_3} \right) \right], \quad (1.4.7)$$

where q_i 's are the cardinal directions and

$$h_i \equiv \left| \frac{\partial \mathbf{r}}{\partial q_i} \right|. \quad (1.4.8)$$

So for spherical coordinates we have

$$\mathbf{r} = (r \cos \phi \sin \theta, r \sin \phi \sin \theta, r \cos \theta), \quad (1.4.9)$$

so

$$h_r = 1 \quad (1.4.10a)$$

$$h_\phi = r \sin \theta \quad (1.4.10b)$$

$$h_\theta = r. \quad (1.4.10c)$$

Substitution into Equation 1.4.7 yields the desired expression. \square

We now write out the time-independent Schrödinger equation in full:

$$-\frac{\hbar^2}{2m} \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \left(\frac{\partial^2 \psi}{\partial \phi^2} \right) \right] + V\psi = E\psi \quad (1.4.11)$$

It is clear that we can use separation of variables to write

$$\psi(r, \phi, \theta) = R(r)Y(\theta, \phi) = R(r)\Theta(\theta)\Phi(\phi). \quad (1.4.12)$$

Using the first equality first, we get after some rearrangement

$$\left\{ \frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) - \frac{2mr^2}{\hbar^2} (V - E) \right\} + \frac{1}{Y} \left\{ \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 Y}{\partial \phi^2} \right\} = 0. \quad (1.4.13)$$

The top bracket involves the potential V and only depends on r , and cannot be solved until the potential has been specified, as we will do in Section 1.4.2. The bottom bracket can be solved as it is, and now we introduce the separation constant $l(l+1)$, where l is in general complex as of now (it will be restricted to integers in the course of solving the equation).

1.4.1.2 The angular equation

The angular equation is

$$\sin \theta \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y}{\partial \theta} \right) + \frac{\partial^2 Y}{\partial \phi^2} = -l(l+1) \sin^2 \theta Y. \quad (1.4.14)$$

Separating further according to Equation 1.4.12 we have

$$\left\{ \frac{1}{\Theta} \left[\sin \theta \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) \right] + l(l+1) \sin^2 \theta \right\} + \frac{1}{\Phi} \frac{d^2 \Phi}{d\phi^2} = 0. \quad (1.4.15)$$

Now introducing the separation constant m^2 , where m is again as of now generally complex, we have

$$\frac{1}{\Theta} \left[\sin \theta \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) \right] + l(l+1) \sin^2 \theta = m^2, \quad (1.4.16)$$

the **polar equation** and

$$\frac{1}{\Phi} \frac{d^2 \Phi}{d\phi^2} = -m^2, \quad (1.4.17)$$

the **azimuthal equation**.

The azimuthal equation is readily solved to give

$$\Phi(\phi) = e^{im\phi}, \quad (1.4.18)$$

where the arbitrary constant is absorbed into Θ and m runs to negative. For the wavefunction to be physical we impose the condition that

$$\Phi(\phi + 2\pi) = \Phi(\phi), \quad (1.4.19)$$

so

$$e^{im(\phi+2\pi)} = e^{im\phi} \Rightarrow e^{2\pi im} = 1 \Rightarrow m = 0, \pm 1, \pm 2, \dots \quad (1.4.20)$$

We have already quantized m to all integers. The normalisation constant is simply $1/\sqrt{2\pi}$. We now solve the polar equation. First of all we introduce the following transformation

$$x \equiv \cos \theta \quad (1.4.21a)$$

$$\Theta(\theta) \equiv P(x), \quad (1.4.21b)$$

with $0 \leq \theta \leq \pi$ and $-1 \leq x \leq 1$, so we have

$$\frac{dx}{d\theta} = -\sin \theta \quad (1.4.22a)$$

$$\sin^2 \theta = 1 - x^2. \quad (1.4.22b)$$

We are all set to go:

$$\frac{1}{\Theta} \left[\sin \theta \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) \right] + l(l+1) \sin^2 \theta = m^2 \quad (1.4.23a)$$

$$\frac{\sin \theta}{P(x)} \frac{\partial x}{\partial \theta} \frac{\partial}{\partial x} \left(\sin \theta \frac{\partial x}{\partial \theta} \frac{\partial}{\partial x} P(x) \right) + l(l+1) \sin^2 \theta = m^2 \quad (1.4.23b)$$

$$\frac{-\sin^2 \theta}{P(x)} \frac{\partial}{\partial x} \left(-\sin^2 \theta \frac{\partial}{\partial x} P(x) \right) + l(l+1) \sin^2 \theta = m^2 \quad (1.4.23c)$$

$$-\frac{\partial}{\partial x} \left[(x^2 - 1) \frac{\partial}{\partial x} P(x) \right] + l(l+1)P(x) - \frac{m^2}{1-x^2} P(x) = 0 \quad (1.4.23d)$$

$$(1-x^2) \frac{d^2}{dx^2} P(x) - 2x \frac{d}{dx} P(x) + \underbrace{\left[l(l+1) - \frac{m^2}{1-x^2} \right]}_{\equiv K} P(x) = 0. \quad (1.4.23e)$$

This is the **general Legendre's equation**. To solve it we must first solve the **Legendre's equation**, where $m = 0$:

$$(1-x^2) \frac{d^2}{dx^2} P(x) - 2x \frac{d}{dx} P(x) + \underbrace{l(l+1)}_{\equiv \beta} P(x) = 0. \quad (1.4.24)$$

As x is inherently dimensionless we can skip the non-dimensionalisation, and also a glance at the equation does show any interesting behaviour at $|x| \rightarrow \infty$, so we can also skip the asymptotic analysis, and jump straight to Frobenius method:

Let

$$P(x) = \sum_{j=0}^{\infty} a_j, \quad (1.4.25)$$

substitution into Equation 1.4.24 gives

$$(1 - x^2) \sum_{j=0}^{\infty} j(j-1)a_j x^{j-2} - 2x \sum_{j=0}^{\infty} j a_j x^{j-1} + \sum_{j=0}^{\infty} a_j \beta x^j = 0 \quad (1.4.26a)$$

$$\sum_{j=0}^{\infty} j(j-1)a_j x^{j-2} - \sum_{j=0}^{\infty} j(j-1)a_j x^j - \sum_{j=0}^{\infty} 2j a_j x^j + \sum_{j=0}^{\infty} a_j \beta x^j = 0 \quad (1.4.26b)$$

$$\sum_{j=2}^{\infty} j(j-1)a_j x^{j-2} - \sum_{j=0}^{\infty} j(j-1)a_j x^j - \sum_{j=0}^{\infty} 2j a_j x^j + \sum_{j=0}^{\infty} a_j \beta x^j = 0 \quad (1.4.26c)$$

$$\sum_{j=0}^{\infty} (j+2)(j+1)a_{j+2} x^j - \sum_{j=0}^{\infty} j(j-1)a_j x^j - \sum_{j=0}^{\infty} 2j a_j x^j + \sum_{j=0}^{\infty} a_j \beta x^j = 0 \quad (1.4.26d)$$

$$\sum_{j=0}^{\infty} [(j+2)(j+1)a_{j+2} - j(j-1)a_j - 2j a_j + a_j \beta] x^j = 0. \quad (1.4.26e)$$

Now let's pause and think what we've done so far: we have set all the powers of x to j , as required by the Frobenius method, and now we have in our hands a $\{a_{j+2}, a\}$ recurrence relation:

$$a_{j+2} = \frac{j(j+1) - l(l+1)}{(j+2)(j+1)} a_j. \quad (1.4.27)$$

We do the same things as we've done with the Hermite polynomials: fix a_0 and a_1 alternately to 0 and terminate series to ensure normalisability. The values are fixed by convention:

$$P_l(1) = 1. \quad (1.4.28)$$

And we observe that we have quantised l to *nonnegative integers, i.e.*,

$$l = 0, 1, 2, \dots. \quad (1.4.29)$$

$P_l(x)$ is called the **Legendre polynomial** and is defined as follows:

Definition 1.4.1.1 (Legendre polynomials). Legendre polynomials are defined by the generating function $T(x, t)$:

$$T(x, t) = (1 - 2xt + t^2)^{-1/2} \equiv \sum_{l=0}^{\infty} P_l(x) t^l, \quad (1.4.30)$$

where $P_l(x)$, the Legendre polynomial of order l is the coefficient of the Taylor series in powers of t . This definition is cumbersome however, and we could alternatively use the Rodrigues' formula:

$$P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2 - 1)^l. \quad (1.4.31)$$

The recurrence relations are

$$(1 - x^2)P'_l = -l x P_l + l P_{l-1} \quad (1.4.32a)$$

$$(l + 1)P_{l+1} = (2l + 1)x P_l - l P_{l-1}. \quad (1.4.32b)$$

todo: supply proof, priority: low

We list the first few Legendre polynomials:

$$P_0 = 1 \quad (1.4.33a)$$

$$P_1 = x \quad (1.4.33b)$$

$$P_2 = \frac{1}{2}(3x^2 - 1) \quad (1.4.33c)$$

$$P_3 = \frac{1}{2}(5x^3 - 3x) \quad (1.4.33d)$$

$$P_4 = \frac{1}{8}(35x^4 - 30x^2 + 3) \quad (1.4.33e)$$

And now we solve the general Legendre's equation. We start by differentiating the Legendre's equation

$$\underbrace{(1 - x^2)P''_l(x)}_{(1)} - \underbrace{2xP'_l(x)}_{(2)} + \underbrace{l(l + 1)P_l(x)}_{(3)} = 0 \quad (1.4.34)$$

m times using the Leibniz formula:

$$\underbrace{(1 - x^2)P_l^{(m+2)} - 2mxP_l^{(m+1)} - m(m + 1)P_l^{(m)}}_{(1)} - \underbrace{[2xP_l^{(m+1)} + 2mP_l^{(m)}]}_{(2)} + \underbrace{l(l + 1)P_l^{(m)}}_{(3)}, \quad (1.4.35)$$

collecting terms we have

$$(1 - x^2)P_l^{(m+2)} - 2x(m + 1)P_l^{(m+1)} + [l(l + 1) - m^2 - m]P_l^{(m)} = 0. \quad (1.4.36)$$

Making the substitution ²

$$P_l^{(m)} \equiv u \equiv v(1-x^2)^{-m/2}, \quad (1.4.37)$$

we obtain

$$\begin{aligned} u' &= v'(1-x^2)^{-m/2} + mvx(1-x^2)^{-m/2-1} \\ &= \left(v' + \frac{mxv}{1-x^2} \right) (1-x^2)^{-m/2}, \end{aligned} \quad (1.4.38)$$

and

$$\begin{aligned} u'' &= \left[v'' + \frac{mv'}{1-x^2} + \frac{2mx^2v}{(1-x^2)^2} \right] (1-x^2)^{-m/2} + \left(v' + \frac{mxv}{1-x^2} \right) (1-x^2)^{-m/2-1} \\ &= \left[v'' + \frac{2mxv'}{1-x^2} + \frac{mv}{1-x^2} + \frac{m(m+2)x^2v}{(1-x^2)^2} \right] (1-x^2)^{-m/2}. \end{aligned} \quad (1.4.39)$$

Substitution into Equation 1.4.35 gives

$$0 = (1-x^2)v'' + 2mxv' + mv + \frac{m(m+2)x^2v}{1-x^2} - 2x(m+1) \left(v' + \frac{mxv}{1-x^2} \right) \quad (1.4.40a)$$

$$\begin{aligned} &+ [l(l+1) - m^2 - m]v \\ &= (1-x^2)v'' + [2mx - 2x(m+1)]v' \\ &+ \left[\frac{2mx^2}{1-x^2} - \frac{2x^2m(m+1)}{1-x^2} + l(l+1) - m^2 \right] v \end{aligned} \quad (1.4.40b)$$

$$= (1-x^2)v'' - 2xv' - m^2v + l(l+1)v + \frac{v}{1-x^2} [m(m+2)x^2 - 2mx^2(m+1)] \quad (1.4.40c)$$

$$= (1-x^2)v'' - 2xv' - m^2 \frac{1-x^2}{1-x^2} v + l(l+1)v - \frac{m^2x^2v}{1-x^2} \quad (1.4.40d)$$

$$= (1-x^2)v'' - 2xv' + \left[l(l+1) - \frac{m^2}{1-x^2} \right] v, \quad (1.4.40e)$$

which is the general Legendre equation. So we now know, though tortuously, that the **associated Legendre polynomials**, defined as

Definition 1.4.1.2 (Associated Legendre polynomials).

$$v_{ml}(x) \equiv P_l^{(m)}(x) = (1-x^2)^{|m|/2} \frac{d^{|m|}}{dx^{|m|}} P_l(x), \quad (1.4.41)$$

are solutions to the general Legendre's equation. Only the magnitude of m is relevant here³ since inspecting the original equation (Equation 1.4.23e) shows that the equation is

²This substitution is introduced here *ad hoc*, but however it arises from the application of Fuch's theorem to Frobenius method as explained in [18] and worked out in [10].

³In some sources such as [19], the polynomials for negative m 's have an appended factor in front, but this is unnecessary for us as the solutions are eventually normalised anyway.

invariant under a sign change of m , so we have our condition

$$|m| \leq l, \Rightarrow m = -l, -l+1, \dots, 0, \dots, +l-1, +l. \quad (1.4.42)$$

This condition arises again, and more naturally, when we study angular momentum. Note that although $P_l^m(x)$ is only a polynomial in x when m is even, $P_l^m(\cos \theta)$ is always a polynomial in $\cos \theta$, multiplied by $\sin \theta$ when m is odd.

At this point we assert that the normalisation constant is

$$N_{lm} = \left[\frac{2l+1}{2} \frac{(l-|m|)!}{(l+|m|)!} \right]^{1/2} \quad (1.4.43)$$

Therefore the normalised **angular wavefunction** is

$$Y_l^m(\theta, \phi) = \epsilon \left[\frac{2l+1}{4\pi} \frac{(l-|m|)!}{(l+|m|)!} \right]^{1/2} P_l^{|m|}(\cos \theta) e^{im\phi}, \quad (1.4.44)$$

where

$$\epsilon = \begin{cases} (-1)^m & \text{for } m > 0; \\ 1 & \text{for } m \leq 0. \end{cases} \quad (1.4.45)$$

This factor is known as the **Condon-Shortley phase**. This sign convention will prove useful when we study the ladder operators for angular momentum: see discussion below Theorem 1.4.3.2. We also notice that the functional makeup of spherical harmonics is invariant under a change of sign.

Definition 1.4.1.3 (Spherical harmonics).

The normalised angular wavefunctions $Y_l^m(\theta, \phi)$ are known as **spherical harmonics**.

Property 1.4.1.3 (Orthogonality of spherical harmonics). Spherical harmonics are orthogonal:

$$\int_0^{2\pi} \int_0^\pi [Y_l^m(\theta, \phi)]^* [Y_{l'}^{m'}(\theta, \phi)] \sin \theta \, d\theta \, d\phi = \delta_{ll'} \delta_{mm'}. \quad (1.4.46)$$

Proof. We prove this by explicitly integrating by parts. **todo: finish this**

□

1.4.1.3 The radial equation

The radial equation is given as

$$\frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) - \frac{2mr^2}{\hbar^2} (V - E)R = l(l+1)R. \quad (1.4.47)$$

We introduce a change of variables:

$$u(r) \equiv rR(r), \quad (1.4.48)$$

such that

$$R = \frac{u}{r}, \quad \frac{dR}{dr} = \frac{1}{r^2} \left(r \frac{du}{dr} - u \right), \quad \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) = r \frac{d^2u}{dr^2}. \quad (1.4.49)$$

So the equation becomes

$$-\frac{\hbar^2}{2m} \frac{d^2u}{dr^2} + \left[V + \frac{\hbar^2}{2m} \frac{l(l+1)}{r^2} \right] u = Eu. \quad (1.4.50)$$

This is identical in form to the one-dimensional Schrödinger equation except for the effective potential:

$$V_{\text{eff}} = V + \frac{\hbar^2}{2m} \frac{l(l+1)}{r^2}, \quad (1.4.51)$$

which contains a second, **centrifugal term**. We can proceed no further until a specific potential $V(r)$ is given.

Infinite spherical well

Consider the potential

$$V(r) = \begin{cases} 0, & \text{if } r \leq a; \\ \infty, & \text{if } r > a. \end{cases} \quad (1.4.52)$$

The wavefunction vanishes outside the well of course, and in the well the Schrödinger equation is

$$\frac{d^2u}{dr^2} = \left[\frac{l(l+1)}{r^2} - k^2 \right] u, \quad \text{where } k \equiv \frac{\sqrt{2mE}}{\hbar} \quad (1.4.53)$$

We solve the case of $l = 0$ first:

$$\frac{d^2u}{dr^2} = -k^2u \Rightarrow u(r) = A \sin(kr) + B \cos(kr), \quad (1.4.54)$$

or

$$\psi(r) = \frac{A \sin(kr)}{r} + \frac{B \cos(kr)}{r}, \quad (1.4.55)$$

$\cos(x)/x$ is not finite at origin hence non-normalisable (**todo: irregular singular point? see Griffiths p.141 footnote 10, priority: low/when going through maths notes**), so we must have $B = 0$, the boundary condition

$$\sin(ka) = 0 \Rightarrow ka = n\pi, \quad n = 1, 2, 3, \dots, \quad (1.4.56)$$

and so

$$E_{n0} = \frac{n^2 \pi^2 \hbar^2}{2ma^2}, \quad (1.4.57)$$

the exact same energy as allowed by the one-dimensional infinite square well in Equation 1.2.31, but however energy in this case is also indexed by l . The full wavefunction for $l = (m =) 0$ is

$$\psi_{n00}(r, \theta, \phi) = AY_0^0(\theta, \phi) \frac{\sin(n\pi r/a)}{r} = \frac{1}{\sqrt{2\pi a}} \frac{\sin(n\pi r/a)}{r}. \quad (1.4.58)$$

The general solution for states with $l \neq 0$ is **todo: outline solution**

$$u(r) = Arj_l(kr) + Brn_l(kr), \quad (1.4.59)$$

where $j_l(x)$ is the **spherical Bessel function** of order l given by

$$j_l(x) \equiv (-x)^l \left(\frac{1}{x} \frac{d}{dx} \right)^l \frac{\sin x}{x}, \quad (1.4.60)$$

and $n_l(x)$ is the **spherical Neumann function** of order l given by

$$n_l(x) \equiv -(-x)^l \left(\frac{1}{x} \frac{d}{dx} \right)^l \frac{\cos x}{x}. \quad (1.4.61)$$

As $\sin x/x$ is finite at origin, so are its derivatives, and $\cos x/x$ is not. So we must again have $B = 0$ and only keep the spherical Bessel functions, so

$$R(r) = Aj_k(kr). \quad (1.4.62)$$

The boundary condition

$$R(a) = 0 \Rightarrow j_l(ka) = 0, \quad (1.4.63)$$

means k is given by zeroes of l -th order spherical Bessel functions, which involves numer-

ical calculations because no neat analytical methods exist. In general, we write

$$k = \frac{1}{a}\beta_{nl}, \quad (1.4.64)$$

where β_{nl} is the n -th zero of the l -th order spherical Bessel function. The energy follows:

$$E_{nl} = \frac{\hbar^2}{2ma^2}\beta_{nl}^2, \quad (1.4.65)$$

and the wavefunctions are

$$\psi_{nlm}(r, \theta, \phi) = A_{nl} j_l(\beta_{nl}r/a) Y_l^m(\theta, \phi), \quad (1.4.66)$$

with each energy level being $(2l + 1)$ -fold degenerate from m .

1.4.2 The hydrogen atom

1.4.2.1 Setting up the radial equation

The potential in a hydrogen atom experienced by the electron is the **Coulombic potential**:

$$V(r) = -\frac{e^2}{4\pi\epsilon_0} \frac{1}{r}. \quad (1.4.67)$$

The Schrödinger equation is

$$-\frac{\hbar^2}{2m} \frac{d^2u}{dr^2} + \left[-\frac{e^2}{4\pi\epsilon_0} \frac{1}{r} + \frac{\hbar^2}{2m} \frac{l(l+1)}{r^2} \right] u = Eu. \quad (1.4.68)$$

Confining ourselves to *bound states*, i.e., $E < 0$, we introduce

$$\kappa \equiv \frac{\sqrt{-2mE}}{\hbar}. \quad (1.4.69)$$

Dividing Equation 1.4.68 by E we have

$$\frac{1}{\kappa^2} \frac{d^2u}{dr^2} = \left[1 - \frac{me^2}{2\pi\epsilon_0\hbar^2\kappa} \frac{1}{r} + \frac{l(l+1)}{(\kappa r)^2} \right] u, \quad (1.4.70)$$

to further simplify, we introduce

$$\rho = \kappa r, \text{ and } \rho_0 \equiv \frac{me^2}{2\pi\epsilon_0\hbar^2\kappa}, \quad (1.4.71)$$

so that

$$\frac{d^2 u}{d\rho^2} = \left[1 - \frac{\rho_0}{\rho} + \frac{l(l+1)}{\rho^2} \right] u. \quad (1.4.72)$$

We have our radial equation.

1.4.2.2 Solving the radial equation

Asymptotic analysis

As $\rho \rightarrow \infty$ (remember we are dealing with r here so $\rho \geq 0$), the radial equation becomes

$$\frac{d^2 u}{d\rho^2} = u. \quad (1.4.73)$$

Taking the physical $e^{-\rho}$ solution only we have

$$u(\rho) \approx A e^{-\rho} \text{ as } \rho \rightarrow \infty. \quad (1.4.74)$$

As $\rho \rightarrow 0$, the $1/\rho^2$ term dominates, so

$$\frac{d^2 u}{d\rho^2} = \frac{l(l+1)}{\rho^2} u, \quad (1.4.75)$$

we posit that the solution must be a polynomial, we try

$$u(\rho) = \sum_{j=0}^{\infty} a_j \rho^j, \quad (1.4.76)$$

substituting back and rearranging we have

$$\sum_{j=0}^{\infty} [j(j-1) - l(l+1)] a_j \rho^{j-2} = 0, \quad (1.4.77)$$

so we have that $j = l+1$ or $j = -l$, giving

$$u(\rho) = C \rho^{l+1} + D \rho^{-l}. \quad (1.4.78)$$

The ρ^{-l} blows up as $\rho \rightarrow 0$, so $D = 0$, so

$$u(\rho) \approx C \rho^{l+1} \text{ as } \rho \rightarrow 0. \quad (1.4.79)$$

This suggest that we write

$$u(\rho) = \rho^{l+1} e^{-\rho} v(\rho). \quad (1.4.80)$$

The derivatives are as follows:

$$\frac{du}{d\rho} = \rho^l e^{-\rho} \left[(l+1-\rho)v + \rho \frac{dv}{d\rho} \right] \quad (1.4.81a)$$

$$\frac{d^2u}{d\rho^2} = \rho^l e^{-\rho} \left\{ \left[-2l-2+\rho + \frac{l(l+1)}{\rho} \right] v + 2(l+1-\rho) \frac{dv}{d\rho} + \rho \frac{d^2v}{d\rho^2} \right\}. \quad (1.4.81b)$$

Now in terms of $v(\rho)$ the radial equation becomes, after a bit of algebra,

$$\rho \frac{d^2v}{d\rho^2} + 2(l+1-\rho) \frac{dv}{d\rho} + [\rho_0 - 2(l+1)]v = 0. \quad (1.4.82)$$

Frobenius method

We adopt the power series solution for $v(\rho)$:

$$v(\rho) = \sum_{j=0}^{\infty} c_j \rho^j. \quad (1.4.83)$$

The required derivatives are

$$\frac{dv}{d\rho} = \sum_{j=0}^{\infty} j c_j \rho^{j-1} = \sum_{j=0}^{\infty} (j+1) c_{j+1} \rho^j \quad (1.4.84a)$$

$$\frac{d^2v}{d\rho^2} = \sum_{j=0}^{\infty} j(j+1) c_{j+1} \rho^{j-1}. \quad (1.4.84b)$$

The equation becomes

$$\begin{aligned} & \sum_{j=0}^{\infty} j(j+1) c_{j+1} \rho^j + 2(l+1) \sum_{j=0}^{\infty} (j+1) c_{j+1} \rho^j \\ & - \sum_{j=0}^{\infty} j c_j \rho^j + [\rho_0 - 2(l+1)] \sum_{j=0}^{\infty} c_j \rho^j = 0. \end{aligned} \quad (1.4.85)$$

This implies

$$\begin{aligned} j(j+1)c_{j+1} + 2(l+1)(j+1)c_{j+1} - 2jc_j + [\rho_0 - 2(l+1)]c_j &= 0 \\ \Rightarrow c_{j+1} &= \frac{2(j+l+1) - \rho_0}{(j+1)(j+2l+2)} c_j. \end{aligned} \quad (1.4.86)$$

Truncation

At large j , hence large ρ , where higher powers dominate, we have

$$c_{j+1} \approx \frac{2j}{j(j+1)} c_j = \frac{2}{j+1} c_j, \quad (1.4.87)$$

so,

$$c_j \approx \frac{2^j}{j!} c_0, \quad (1.4.88)$$

then

$$v(\rho) \approx c_0 \sum_{j=0}^{\infty} \frac{2^j}{j!} \rho^j = c_0 e^{2\rho}. \quad (1.4.89)$$

Therefore

$$u(\rho) = c_0 \rho^{l+1} e^{\rho}, \quad (1.4.90)$$

which is *bad* behaviour. So we must terminate the series, with some maximal integer j_{\max} such that

$$c_{(j_{\max}+1)} = 0, \quad (1.4.91)$$

which implies

$$2 \underbrace{(j_{\max} + l + 1)}_{\equiv n} - \rho_0 = 0. \quad (1.4.92)$$

We now must have

$$\rho_{0,n} = 2n. \quad (1.4.93)$$

With this we have quantised our energy (which ρ_0 depends on). We can conclude about the degeneracy of the system:

Property 1.4.2.1 (Degeneracy of hydrogen atom). From Equation 1.4.92, we can see that, given an arbitrary n , with j_{\max} as a freely varying integer, we have $(n-1)$ possible values of l :

$$l = 0, 1, 2, \dots, n-1. \quad (1.4.94)$$

And for each value of l , we refer back to Equation 1.4.42 to see that there are $(2l+1)$ possible values of m . The total degeneracy of E_n is therefore

$$d(n) = \sum_{l=0}^{n-1} (2l+1) = n^2. \quad (1.4.95)$$

Additionally, we make the following definitions:

Definition 1.4.2.1 (Bohr formula and Bohr radius). Remembering that

$$\kappa = \frac{me^2}{2\pi\epsilon_0\hbar^2\rho_0}, \quad (1.4.96)$$

we can write

$$E_n = -\frac{\hbar^2\kappa_n^2}{2m} = -\frac{me^4}{8\pi^2\epsilon_0^2\hbar^2\rho_{0,n}^2} = -\left[\frac{m}{2\hbar^2}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2\right]\frac{1}{n^2} \equiv \frac{E_1}{n^2}, \quad (1.4.97)$$

where n can be positive integers.

Also we have

$$\kappa = \left(\frac{me^2}{4\pi\epsilon_0\hbar^2}\right)\frac{1}{n} \equiv \frac{1}{an}, \quad (1.4.98)$$

where a is the **Bohr radius**, and $E_1 \equiv -1 \text{ Ry} \equiv -1/2 \text{ Ha}$.

Now we should return to the solution to the radial equation:

$$R_{nl}(r) = \frac{1}{r}u_{nl}(\rho) = \frac{1}{r}\rho^{l+1}e^{-\rho}v(\rho), \quad (1.4.99)$$

where $v_{nl}(\rho)$ is a polynomial determined (up to an overall normalisation factor) by the recursion formula

$$c_{j+1} = \frac{2(j+l+1-n)}{(j+1)(j+2l+2)}c_j. \quad (1.4.100)$$

The degree of the polynomial is then $j_{\max} = n - l - 1$ in ρ . The polynomial can be *written in terms of associated Laguerre polynomials*:

$$v(\rho) = L_{n-l-1}^{2l+1}(2\rho). \quad (1.4.101)$$

They are defined as follows

Definition 1.4.2.2 (Laguerre polynomials).

$$L_q(x) \equiv e^x \frac{d^q}{dx^q}(e^{-x}x^q) \quad (1.4.102)$$

is the Laguerre polynomial of degree q , and

$$L_{q-p}^p(x) \equiv (-1)^p \frac{d^p}{dx^p}[L_q(x)], \quad (1.4.103)$$

is the associated Laguerre polynomial.

Now the complete and normalised radial solution can be written as

$$\begin{aligned}
 R_{nl}(r) &= A_0 \cdot \frac{1}{na} \left(\frac{r}{na} \right)^l e^{-r/na} \left[L_{n-l-1}^{2l+1}(2r/na) \right] \\
 &= \left[\left(\frac{2}{na} \right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3} \right]^{1/2} e^{-r/na} \left(\frac{2r}{na} \right)^l \left[L_{n-l-1}^{2l+1}(2r/na) \right].
 \end{aligned} \tag{1.4.104}$$

Property 1.4.2.2 (Complete hydrogen wavefunction). The complete solution to this eigenvalue problem

$$H\psi = E\psi \tag{1.4.105}$$

is

$$\begin{aligned}
 \psi_{nlm} &= R_{nl}(r) Y_l^m(\theta, \phi) \\
 &= \epsilon \left[\left(\frac{2}{na} \right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3} \frac{2l+1}{4\pi} \frac{(l-|m|)!}{(l+|m|)!} \right]^{1/2} \\
 &\quad \times e^{-r/na} \left(\frac{2r}{na} \right)^l L_{n-l-1}^{2l+1}(2r/na) P_l^m(\cos \theta) e^{im\phi},
 \end{aligned} \tag{1.4.106}$$

with

$$E = \frac{E_1}{n^2}, \quad n = 1, 2, 3, \dots, \text{ or, in atomic units (Hartree), } E = \frac{1}{2n^2}, \tag{1.4.107}$$

where $E_1 \approx -13.6$ eV is the ground state energy of hydrogen.

The three indices specifying the wavefunction is called the **principal, azimuthal and magnetic quantum numbers**, respectively for n , l (henceforce styled as ℓ to avoid confusion) and m .

1.4.2.3 Larger nuclei

In many subsequent treatments of heavier atoms, we can use the form of hydrogen wavefunctions as a first approximation to their true wavefunctions. And this requires us to use the potential of the form

$$V(r) = -\frac{e^2}{4\pi\epsilon_0} \frac{Z}{r}. \tag{1.4.108}$$

As the constant e^2 appear nowhere else, we just need to replace e^2 with Ze^2 everywhere. So now, as the full wavefunction involves it, the Bohr radius is replaced with

$$a \leftrightarrow \frac{a}{Z}. \tag{1.4.109}$$

So the full wavefunction is

$$R_{n\ell}(r) = - \left[\left(\frac{2}{na} \right)^3 \frac{(n-\ell-1)!}{2n[(n+\ell)^3]} \right]^{1/2} Z^{\ell+3/2} e^{-Zr/na} \left(\frac{2r}{na} \right)^\ell [L_{n-\ell-1}^{2\ell+1}(2Zr/na)], \quad (1.4.110)$$

with the energy

$$E_n = -\frac{Z^2 E_1}{n^2} \quad (1.4.111)$$

1.4.2.4 Discussion of hydrogen atomic orbitals

Constructing real orbitals

By conventions in chemistry the azimuthal quantum numbers ℓ are given names s, p, d, f, \dots for $\ell = 0, 1, 2, 3, \dots$. Now if we list the $2p$ wavefunctions:

$$\begin{aligned} \psi_{2p0} &= \left(\frac{1}{32\pi a^5} \right)^{1/2} r e^{-r/2a} \cos \theta \\ \psi_{2p\pm 1} &= \left(\frac{1}{64\pi a^5} \right)^{1/2} r e^{-r/2a} \sin \theta e^{\pm i\phi}. \end{aligned} \quad (1.4.112)$$

Clearly, when $m \neq 0$, the wavefunctions won't be real, and we again invoke the all powerful Theorem 1.2.1.3 - we can always take linear combinations of them to make them real, as long as *they have the same energy*, which our wavefunctions do (m gives the $2\ell + 1$ degenerate states). Now, we tend to visualise better with Cartesian coordinates and that's what we'll re-write our wavefunctions with:

$$\begin{aligned} \psi_{2p,0} &\propto z e^{-r/2a} \\ \psi_{2p,+1} &\propto r e^{-r/2a} \sin \theta e^{i\phi} = r \sin \theta (\cos \phi + i \sin \phi) e^{-r/2a} = (x + iy) e^{-r/2a} \\ \psi_{2p,-1} &\propto (x - iy) e^{-r/2a}. \end{aligned} \quad (1.4.113)$$

Following the recipe given in the proof of Theorem 1.2.1.3 (noting that $\psi_{2p,+1}$ and $\psi_{2p,-1}$ are complex conjugates of each other) and by convention (which is just so that the plots actually align with x and y axes), we construct

Definition 1.4.2.3 (Construction of p orbitals).

$$\psi_{2p_x} \equiv \frac{1}{\sqrt{2}}(-\psi_{2p,+1} + \psi_{2p,-1}) \propto x e^{-r/2a} \quad (1.4.114a)$$

$$\psi_{2p_y} \equiv \frac{1}{\sqrt{2}}i(\psi_{2p,+1} + \psi_{2p,-1}) \propto y e^{-r/2a} \quad (1.4.114b)$$

$$\psi_{2p_z} \equiv \psi_{2p,0} \propto z e^{-r/2a}. \quad (1.4.114c)$$

The factors of $1/\sqrt{2}$ are only there to normalise the combined wavefunction and plays no role whatsoever for Theorem 1.2.1.3 to apply.

Now note that Theorem 1.2.1.3 only applies when the eigenvalues are the same, which these are, for H and L^2 but not for L_z (see Theorem 1.4.3.2), so the linearly combined orbitals p_x and p_y **todo-supo: should not be acted on by L_z ?**

Now, for $3d$ orbitals, we have

$$\begin{aligned} \psi_{3d,0} &= \frac{1}{81\sqrt{6\pi a^7}} r^2 e^{-r/3a} (3 \cos^2 \theta - 1) \\ \psi_{3d,\pm 1} &= \frac{1}{81\sqrt{\pi a^7}} r^2 e^{-r/3a} \sin \theta \cos \theta e^{\pm i\phi} \\ \psi_{3d,\pm 2} &= \frac{1}{162\sqrt{\pi a^7}} r^2 e^{-r/3a} \sin^2 \theta e^{\pm 2i\phi}. \end{aligned} \quad (1.4.115)$$

Re-writing these in Cartesian we have

$$\psi_{3d,0} \propto (3z^2 - r^2) e^{-r/3a} \quad (1.4.116a)$$

$$\psi_{3d,\pm 1} \propto z(x \pm iy) e^{-r/3a} \quad (1.4.116b)$$

$$\psi_{3d,\pm 2} \propto (x^2 - y^2 \pm ixy) e^{-r/3a}. \quad (1.4.116c)$$

Definition 1.4.2.4 (Construction of d orbitals). Making linear combinations for same $|m|$'s, we have

$$\psi_{3d_{z^2}} = \psi_{3d,0} \propto (3z^2 - r^2) e^{-r/3a} \quad (1.4.117a)$$

$$\psi_{3d_{xz}} = \frac{1}{\sqrt{2}}(\psi_{3d,+1} + \psi_{3d,-1}) \propto xye^{-r/3a} \quad (1.4.117b)$$

$$\psi_{3d_{yz}} = \frac{1}{\sqrt{2}}i(-\psi_{3d,+1} + \psi_{3d,-1}) \propto yze^{-r/3a} \quad (1.4.117c)$$

$$\psi_{3d_{x^2-y^2}} = \frac{1}{\sqrt{2}}(\psi_{3d,+2} + \psi_{3d,-2}) \propto (x^2 - y^2) e^{-r/3a} \quad (1.4.117d)$$

$$\psi_{3d_{xy}} = \frac{1}{\sqrt{2}}i(-\psi_{3d,+2} + \psi_{3d,-2}) \propto xye^{-r/3a}. \quad (1.4.117e)$$

Nodal structure

There are two types of nodes in complete hydrogen wavefunctions: radial and angular. The **radial nodes** arise entirely from the *associated Laguerre polynomials* as the exponential and the polynomial terms do not introduce nodes. The radial wavefunctions with the same l , i.e., R_{ns} , R_{np} and so on are the solutions to the *same* eigenproblem and so they have to be orthogonal to each other wrt radial integration. This means that each successive radial wavefunction must have one more radial node as n increases by one. Another reason this should be the case is because with more nodes the energy will rise, as is expected for higher n .

The **angular nodes** arise from the *associated Legendre polynomials*, which means that all the nodes are *polar* in nature. The angular equation however is *not* an eigenvalue equation, and as such we do not have *a priori* reason to assert that they are orthogonal. But nonetheless they are indeed orthogonal, and it's proven in Property 1.4.1.3. So we again have the same behaviour as in the radial wavefunction with increasing l giving increasing number of nodes.

In all, the atomic orbitals with principal quantum number n all have $n - 1$ nodal surfaces. Of these, l are angular nodes and the rest are radial.

Radial probability density

The radial probability density is defined by

Definition 1.4.2.5 (Radial probability density).

$$\begin{aligned}
 P_{nl}(r) \, dr &= \int_0^{2\pi} \int_0^\pi |\psi|^2 r^2 \, dr \sin \theta \, d\theta \, d\phi \\
 &= \int_0^{2\pi} \int_0^\pi (R_{nl} Y_{lm})^* R_{nl} Y_{lm} r^2 \, dr \sin \theta \, d\theta \, d\phi \\
 &= r^2 R_{nl}^2 \, dr \underbrace{\int_0^{2\pi} \int_0^\pi Y_{lm}^* Y_{lm} \sin \theta \, d\theta \, d\phi}_{1 \text{ since normalised}} \\
 &= r^2 R_{nl}^2 \, dr
 \end{aligned}
 \tag{1.4.118}$$

We can find the *most probable location* to find an electron by finding the maximum value of the probability distribution (the expression without dr). Note that this is emphatically not the same as finding the expectation of r as mode and mean of a distribution are in general different.

1.4.3 Angular momentum

1.4.3.1 Operator properties

The classical angular momentum is

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}, \quad (1.4.119)$$

which means

$$L_x = yp_z - zp_y \quad (1.4.120)$$

and cyclical permutations thereof. We make the canonical substitutions ($p_x \rightarrow -i\hbar\partial/\partial x$ and so on) to get the quantum mechanical angular momentum operators:

Definition 1.4.3.1 (Angular momentum operators).

$$\begin{aligned} L_x &= -i\hbar \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \\ L_y &= -i\hbar \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) \\ L_z &= -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right). \end{aligned} \quad (1.4.121)$$

The operators have the following commutator relationships:

Property 1.4.3.1 (Fundamental angular momentum commutator relations). The commutators **do not commute with each other**:

$$\begin{aligned} [L_x, L_y] &= [yp_z - zp_y, zp_x - xp_z] \\ &= [yp_z, zp_x] - [yp_z, xp_z] - [zp_y, zp_x] + [zp_y, xp_z] \\ &= yp_x[p_z, z] - 0 - 0 + xp_y[z, p_z] \\ &= i\hbar(xp_y - yp_x) \\ &= i\hbar L_z. \end{aligned} \quad (1.4.122)$$

In the third and fourth lines we used the 3D canonical commutation relations (Property 1.4.1.1). Cyclic permutations follow:

$$[L_x, L_y] = i\hbar L_z, \quad [L_z, L_x] = i\hbar L_y, \quad [L_y, L_z] = i\hbar L_x. \quad (1.4.123)$$

Corollary 1.4.3.1 (Uncertainty). The preceding commutator relation means that the three components of angular momentum are **incompatible observables**, *i.e.*, they cannot be simultaneously measured. We obtain, from the generalised uncertainty principle (Theorem 1.3.5.1):

$$\sigma_{L_x}^2 \sigma_{L_y}^2 \geq \left(\frac{1}{2i} \langle i\hbar L_z \rangle \right)^2 = \frac{\hbar^2}{4} \langle L_z \rangle^2, \quad (1.4.124)$$

which implies

$$\sigma_{L_x} \sigma_{L_y} \geq \frac{\hbar}{2} |\langle L_z \rangle|. \quad (1.4.125)$$

We now introduce the operator for the square of total momentum, L^2 :

$$L^2 \equiv L_x^2 + L_y^2 + L_z^2. \quad (1.4.126)$$

Property 1.4.3.2 (Commutator relations with L^2). We can write that

$$\begin{aligned} [L^2, L_x] &= [L_x^2, L_x] + [L_y^2, L_x] + [L_z^2, L_x] \\ &= L_y[L_y, L_x] + [L_y, L_x]L_y + L_z[L_z, L_x] + [L_z, L_x]L_z \\ &= -i\hbar([L_y, L_z] - [L_z, L_y]) \\ &= 0, \end{aligned} \quad (1.4.127)$$

where in the second equality we have used the associativity of commutators (Property 1.3.5.1). We can compactly write

$$[L^2, \mathbf{L}] = 0. \quad (1.4.128)$$

This means that we can simultaneously measure the scalar total angular momentum and **one** component of the vectorial angular momentum, which is *conventionally chosen* to be L_z .

1.4.3.2 Angular momentum ladder operators

The above property means that L^2 and L_z (conventionally chosen) can have simultaneous eigenfunctions (Theorem 1.3.5.3), which is to say

$$L^2 f = \lambda f \text{ and } L_z f = \mu f. \quad (1.4.129)$$

And again we define two ladder operators:

Definition 1.4.3.2 (Angular momentum ladder operators). Let

$$L_{\pm} \equiv L_x \pm iL_y. \quad (1.4.130)$$

The commutators with the ladder operators will be useful later on:

Property 1.4.3.3 (Commutators with L_{\pm}). The commutator with L_z is

$$[L_z, L_{\pm}] = [L_z, L_x] \pm i[L_z, L_y] = i\hbar L_y \pm i(-i\hbar L_x) = \pm\hbar(L_x \pm iL_y). \quad (1.4.131)$$

And the commutator with L^2 is easy:

$$[L^2, L_{\pm}] = [L^2, L_x] \pm i[L^2, L_y] = 0. \quad (1.4.132)$$

Property 1.4.3.4 (Further property).

$$\begin{aligned} L_{\pm}L_{\mp} &= (L_x \pm iL_y)(L_x \mp iL_y) \\ &= L_x^2 + L_y^2 \mp i(L_xL_y - L_yL_x) \\ &= L^2 - L_z^2 \pm \hbar L_z \end{aligned} \quad (1.4.133)$$

which implies

$$L^2 = L_{\pm}L_{\mp} + L_z^2 \mp \hbar L_z. \quad (1.4.134)$$

From Equation 1.4.132 we can write

$$L^2L_{\pm}f = L_{\pm}L^2f = L_{\pm}\lambda f = \lambda L_{\pm}f. \quad (1.4.135)$$

This means that $L_{\pm}f$ is an eigenfunction of L_{\pm} and L_z , with the same eigenvalue λ , meaning that the total scalar angular momentum is unchanged. And from Equation 1.4.131 we can write

$$\begin{aligned} L_zL_{\pm}f &= (L_zL_{\pm} - L_{\pm}L_z)f + L_{\pm}L_zf \\ &= \pm\hbar L_{\pm}f + L_{\pm}\mu f \\ &= (\mu \pm \hbar)(L_{\pm}f). \end{aligned} \quad (1.4.136)$$

Evidently, L_+ increases the eigenvalue of L_z by \hbar and vice versa for L_- .

So we now know that, given total squared angular momentum λ , we can obtain, by way of our ladder operators, a range of states. But there must be a upper and lower limit of L_z such that it's within 0 to maximum angular momentum.

The ‘top rung’ wavefunction must have that

$$L_+ f_t = 0. \quad (1.4.137)$$

We now let the eigenvalue of L_z be $\hbar l$, where l is *not necessarily* the azimuthal quantum number ℓ as discussed before, but we’ll show in time that it actually is. Anyway, we can write

$$L_z f_t = \hbar l f_t, \quad L^2 f_t = \lambda f_t. \quad (1.4.138)$$

It follows, from Property 1.4.3.4, that

$$L^2 f_t = (L_- L_+ + L_z^2 + \hbar L_z) f_t = (0 + \hbar^2 l^2 + \hbar^2 l) f_t = \hbar^2 l(l+1) f_t, \quad (1.4.139)$$

where we have chosen to use $L_- L_+$ because we need to have L_+ directly left-multiplying f_t .

So we conclude that

$$\lambda = \hbar^2 l(l+1). \quad (1.4.140)$$

For the bottom rung, we must have that

$$L_- f_b = 0. \quad (1.4.141)$$

Let the eigenvalue of L_z at bottom rung be $\hbar l'$

$$L_z f_b = \hbar l' f_b, \quad L^2 f_b = \lambda. \quad (1.4.142)$$

Similarly, we write

$$L^2 f_b = (L_+ L_- + L_z^2 - \hbar L_z) f_b = (0 + \hbar^2 l'^2 - \hbar^2 l') f_b = \hbar^2 l'(l'-1) f_b, \quad (1.4.143)$$

so

$$\lambda = \hbar^2 l'(l'-1). \quad (1.4.144)$$

Because we must have that $l > l'$,

$$l' = -l. \quad (1.4.145)$$

So now we have our range of eigenvalues of L_z , and let’s say the eigenvalues are $m\hbar$, where m again is not yet necessarily the magnetic quantum number, and that m goes from $-l$ to l in some number of integer steps, say N , so

$$-l + N = l \Rightarrow l = \frac{N}{2}. \quad (1.4.146)$$

We have now effectively *quantised* l to integers and half-integers⁴. And given a value of l we have $2l + 1$ values of m .

Property 1.4.3.5 (Proportionality constant of ladder operators). We have

$$L_{\pm} f_l^m = A_l^m f_l^{m\pm 1}, \quad (1.4.147)$$

where

$$A_l^m = \hbar \sqrt{l(l+1) - m^2 \mp m} = \hbar \sqrt{(l \mp m)(l \pm m + 1)}. \quad (1.4.148)$$

Proof. We know that L_x and L_y are hermitian as they are observables (Theorem 1.3.2.1). So we know that L_{\pm} and L_{\mp} are hermitian conjugates:

$$L_{\pm}^{\dagger} = (L_x + iL_y)^{\dagger} = L_x - iL_y = L_{\mp}. \quad (1.4.149)$$

Since we know

$$L_{\pm} f_l^m = A_l^m f_l^{m\pm 1}, \quad (1.4.150)$$

so

$$\begin{aligned} \langle L_{\pm} f_l^m | L_{\pm} f_l^m \rangle &= \langle f_l^m | L_{\mp} L_{\pm} f_l^m \rangle \\ &= \langle f_l^m | (L^2 - L_z^2 \mp \hbar L_z) f_l^m \rangle \\ &= \hbar^2 [l(l+1) - m^2 \mp m] \langle f_l^m | f_l^m \rangle \\ &\equiv (A_l^m)^2 \langle f_l^{m\pm 1} | f_l^{m\pm 1} \rangle. \end{aligned} \quad (1.4.151)$$

As all the eigenfunctions are assumed to be normalised, we must have that

$$A_l^m = \hbar \sqrt{l(l+1) - m^2 \mp m} = \hbar \sqrt{(l \mp m)(l \pm m + 1)}. \quad (1.4.152)$$

Applying L_+ to $f_l = f_l^l$ or applying L_- to $f_l = f_l^{-l}$ gives zero, as required. \square

We prove some further commutator relationships.

⁴Here we have no reason to exclude the half-integers because this is an entirely algebraic treatment. But when we link it to the spherical harmonics in Equations 1.4.169 and 1.4.170, we will have to restrict ourselves to integers. But in the treatment of spin angular momentum, we again have no reason to restrict to integers and it'll turn out that we indeed need to admit half-integers.

Property 1.4.3.6 (Further commutator relations involving L).

$$[L_z, x] = i\hbar y, [L_z, y] = -i\hbar x, [L_z, z] = 0.$$

$$\text{compactly, } [L_{x_k}, x_i] = i\hbar \epsilon_{ijk} x_j \quad (1.4.153)$$

And also

$$[L_{x_k}, p_{x_i}] = i\hbar \epsilon_{ijk} p_{x_j}. \quad (1.4.154)$$

Proof is similar to ones we have gone through before.

We can apply this to evaluate $[L_z, L_x]$:

$$\begin{aligned} [L_z, L_x] &= [L_z, (yp_z - zp_y)] \\ &= [L_z, yp_z] - [L_z, zp_y] \\ &= y[L_z, p_z] + [L_z, y]p_z - z[L_z, p_y] - [L_z, z]p_y \\ &= i\hbar(zp_y - xp_z) \\ &= i\hbar L_y. \end{aligned} \quad (1.4.155)$$

Now we evaluate $[L_z, r^2]$:

$$\begin{aligned} [L_z, r^2]f &= [xp_y - yp_x, r^2]f \\ &= [xp_y, r^2]f - [yp_x, r^2]f \\ &= i\hbar(2xyf - xr^2f_y - 2xyf - yr^2f_x - r^2xf_y + r^2yf_x) \\ &= 0. \end{aligned} \quad (1.4.156)$$

And

$$\begin{aligned} [L_z, p^2]f &= [L_z, p_x^2 + p_y^2]f \\ &= [xp_y - yp_x, p_x^2]f + [xp_y - yp_x, p_y^2]f \\ &= [xp_y, p_x^2]f - [yp_x, p_x^2]f \\ &= i\hbar^3(xf_{xxy} - xf_{xyx} - 2f_{xy} - yf_{yyx} + yf_{yyx} + 2f_{xy}) \\ &= 0 \end{aligned} \quad (1.4.157)$$

And as the Hamiltonian is

$$H \equiv \frac{p^2}{2m} + V(r), \quad (1.4.158)$$

where we have assume that the potential is a function of *only* r . As L_z commutes with p^2 and r^2 , and by extension, *any* function of only r (as r^2 can be measured to any degree of accuracy with L_z , any function of the *scalar* r can be evaluated also). We therefore have an important theorem

Theorem 1.4.3.1 (Commutator of angular momentum and Hamiltonian). H , L_z and L^2 are mutually compatible observables.

1.4.3.3 Eigenfunctions

We need to rewrite L 's in spherical coordinates. We first note that in classical mechanics,

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \quad (1.4.159)$$

making the canonical substitutions we have, for our vectorial quantum mechanical operator,

$$\hat{\mathbf{L}} = \frac{\hbar}{i}(\mathbf{r} \times \nabla) = \frac{\hbar}{i}r(\hat{\mathbf{r}} \times \nabla) \quad (1.4.160a)$$

$$= \frac{\hbar}{i}r\hat{\mathbf{r}} \times \left(\hat{\mathbf{r}} \frac{\partial}{\partial r} + \hat{\boldsymbol{\theta}} \frac{1}{r} \frac{\partial}{\partial \theta} + \hat{\boldsymbol{\phi}} \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} \right) \quad (1.4.160b)$$

$$= \frac{\hbar}{i} \left[r(\hat{\mathbf{r}} \times \hat{\mathbf{r}}) \frac{\partial}{\partial r} + (\hat{\mathbf{r}} \times \hat{\boldsymbol{\theta}}) \frac{\partial}{\partial \theta} + (\hat{\mathbf{r}} \times \hat{\boldsymbol{\phi}}) \frac{1}{\sin \theta} \frac{\partial}{\partial \phi} \right] \quad (1.4.160c)$$

$$= \frac{\hbar}{i} \left(\hat{\boldsymbol{\phi}} \frac{\partial}{\partial \theta} - \hat{\boldsymbol{\theta}} \frac{1}{\sin \theta} \frac{\partial}{\partial \phi} \right) \quad (1.4.160d)$$

$$= \frac{\hbar}{i} \left[(-\sin \phi \hat{\mathbf{i}} + \cos \phi \hat{\mathbf{j}}) \frac{\partial}{\partial \theta} - (\cos \theta \cos \phi \hat{\mathbf{i}} + \cos \theta \sin \phi \hat{\mathbf{j}} - \sin \theta \hat{\mathbf{k}}) \frac{1}{\sin \theta} \frac{\partial}{\partial \phi} \right], \quad (1.4.160e)$$

where in Equation 1.4.160d we have used the fact that the unit vectors are oriented $\{\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}\}$ as related by the right-hand rule. Also, the unit vectors are obtained by

$$\hat{\mathbf{e}}_i = \frac{\partial \mathbf{r} / \partial e_i}{|\partial \mathbf{r} / \partial e_i|}. \quad (1.4.161)$$

We can identify from Equation 1.4.160e the three components:

Property 1.4.3.7 (Angular momentum operators in spherical coordinates).

$$L_x = \frac{\hbar}{i} \left(-\sin \phi \frac{\partial}{\partial \theta} - \cos \phi \cot \theta \frac{\partial}{\partial \phi} \right), \quad (1.4.162a)$$

$$L_y = \frac{\hbar}{i} \left(+\cos \phi \frac{\partial}{\partial \theta} - \sin \phi \cot \theta \frac{\partial}{\partial \phi} \right), \quad (1.4.162b)$$

$$L_z = \frac{\hbar}{i} \frac{\partial}{\partial \phi}. \quad (1.4.162c)$$

The same result can be obtained by chain rule (and it's admittedly faster).

Property 1.4.3.8 (Angular momentum ladder operators in spherical coordinates).

$$L_{\pm} = \pm \hbar e^{\pm i\phi} \left(\frac{\partial}{\partial \theta} \pm i \cot \theta \frac{\partial}{\partial \phi} \right), \quad (1.4.163)$$

$$L_+ L_- = -\hbar^2 \left(\frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \cot^2 \theta \frac{\partial^2}{\partial \phi^2} + i \frac{\partial}{\partial \phi} \right), \quad (1.4.164)$$

$$L^2 = -\hbar^2 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right]. \quad (1.4.165)$$

Proof. We can readily write

$$\begin{aligned} L_{\pm} &= L_x \pm i L_y = \frac{\hbar}{i} \left[(-\sin \phi \pm i \cos \phi) \frac{\partial}{\partial \phi} - (\cos \phi \pm i \sin \phi) \cot \theta \frac{\partial}{\partial \theta} \right] \\ &= \pm \hbar e^{\pm i\phi} \left(\frac{\partial}{\partial \theta} \pm i \cot \theta \frac{\partial}{\partial \phi} \right) \end{aligned} \quad (1.4.166)$$

Now the product of two ladder operators is more arduous:

$$\begin{aligned} L_+ L_- f &= \hbar e^{i\phi} \left(\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right) \left[-\hbar e^{-i\phi} \left(\frac{\partial}{\partial \theta} - i \cot \theta \frac{\partial}{\partial \phi} \right) \right] f \\ &= -\hbar^2 e^{i\phi} \left(\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right) \left[e^{-i\phi} (f_{\theta} - i \cot \theta f_{\phi}) \right] \\ &= -\hbar^2 e^{i\phi} \left[\frac{\partial}{\partial \theta} (e^{i\phi} f_{\theta} - i \cot \theta e^{-i\phi} f_{\phi}) + i \cot \theta \frac{\partial}{\partial \phi} (e^{-i\phi} f_{\theta} - i \cot \theta e^{-i\phi} f_{\phi}) \right] \\ &= \dots \\ &= -\hbar^2 e^{i\phi} (e^{-i\phi} f_{\theta\theta} + i e^{-i\phi} \csc^2 \theta f_{\phi} + \cot \theta e^{-i\phi} f_{\theta} - \cot^2 \theta e^{-i\phi} + \cot^2 \theta e^{-i\phi} f_{\phi\phi}) \\ &= -\hbar^2 \left(\frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \cot^2 \theta \frac{\partial^2}{\partial \phi^2} + i \frac{\partial}{\partial \phi} \right) f. \end{aligned} \quad (1.4.167)$$

And finally, using Equation 1.4.135, we can write

$$\begin{aligned} L^2 &= L_+ L_- + L_z^2 - \hbar L_z \\ &= -\hbar^2 \left(\frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \csc^2 \theta \frac{\partial^2}{\partial \phi^2} \right) \\ &= -\hbar^2 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right]. \end{aligned} \quad (1.4.168)$$

□

We're all set to determine $f_l^m(\theta, \phi)$, which is an eigenfunction of L^2 and L_z , with eigenvalues $\hbar^2 l(l+1)$ and $m\hbar$ respectively, remembering that m and l do not correspond to the spherical harmonics indices *yet*. We can therefore write:

$$L^2 f_l^m = -\hbar^2 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right] f_l^m = \hbar^2 l(l+1) f_l^m, \quad (1.4.169)$$

which is exactly the same as the angular equation (Equation 1.4.15). For L_z we can write

$$L_z f_l^m = \frac{\hbar}{i} \frac{\partial}{\partial \phi} f_l^m = \hbar m f_l^m, \quad (1.4.170)$$

which is the azimuthal equation (Equation 1.4.17). We can therefore identify f_l^m with $Y_\ell^m(\theta, \phi)$ and therefore m with the magnetic quantum number and l with the azimuthal quantum number. We reach the important conclusion

Theorem 1.4.3.2 (Eigenfunctions of L^2 , L_z and H). The spherical harmonic, $Y_\ell^m(\theta, \phi)$ are the eigenfunctions of the three *commuting* operators H , L^2 and L_z with the eigenequations:

$$H\psi = E\psi, \quad L^2\psi = \hbar^2 l(l+1)\psi, \quad L_z\psi = \hbar m\psi. \quad (1.4.171)$$

Incidentally⁵, now that we have determined that the spherical harmonics are the function f here, we can see how the introduction of the Condon-Shortley phase in Equation 1.4.45 simplifies notation: the ladder operators alternate the signs of the wavefunctions by means of differentiation of $\cos \theta$ and so on, so to be able to write

$$L_\pm Y_l^m = A Y_l^{m\pm 1}, \quad (1.4.172)$$

where A is always positive as in Property 1.4.3.5, we have to absorb the phase factor into the spherical harmonics expression.

Another aside is, the Schrödinger equation in spherical coordinates can be more compactly written as

$$\frac{1}{2mr^2} \left[-\hbar^2 \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + L^2 \right] \psi + V\psi = E\psi. \quad (1.4.173)$$

⁵Adapted from p.778 of [1]

Corollary 1.4.3.2 (Spherical harmonics at the top of the ladder). The angular wavefunction at the top of the ladder is

$$Y_\ell^\ell(\theta, \phi) = Ce^{i\ell\phi} \sin^\ell \theta, \quad (1.4.174)$$

where C is the normalisation constant.

Proof. We construct the proof with the help of the ladder operators:

We know that

$$L_+ Y_\ell^\ell = 0 \Rightarrow \hbar e^{i\phi} \left[\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right] \underbrace{Y_\ell^\ell}_{\equiv f} = 0. \quad (1.4.175)$$

We can write

$$f_\theta + i \cot \theta f_\phi = 0 \Rightarrow \tan \theta f_\theta = -i f_\phi, \quad (1.4.176)$$

where we have noted that $\hbar e^{i\phi}$ is always non-zero, hence the content of the square bracket must be zero. We try separation of variables:

$$f(\theta, \phi) = g(\theta)h(\phi). \quad (1.4.177)$$

Substituting we have

$$\tan \theta \frac{g_\theta}{g} = -i \frac{h_\phi}{h} \equiv \ell, \quad (1.4.178)$$

where we made the separation constant ℓ because we *know* the index can only be ℓ .

Solving both equations give us

$$f = Ce^{i\ell\phi} \sin^\ell \theta. \quad (1.4.179)$$

□

1.4.4 The rigid rotor

We finally consider the rigid rotor, which is a model of two particles with mass m_1 and m_2 , at fixed radius r_1 and r_2 away from the centre of mass. There is no V in the Schrödinger equation in this case, so

$$H = K = \frac{1}{2I} L^2. \quad (1.4.180)$$

Now, to calculate moment of inertia, we know

$$\begin{aligned} I &= m_1 r_1^2 + m_2 r_2^2 \\ &= \frac{1}{m_1 + m_2} \left(m_1 m_2 r_1^2 + m_1 m_2 r_2^2 + m_1^2 r_1^2 + m_2^2 r_2^2 \right). \end{aligned} \quad (1.4.181)$$

Now that, we recall that our two masses are rotating around the centre of mass, **defined** by

$$m_1 r_1 = m_2 r_2. \quad (1.4.182)$$

So we have

$$\begin{aligned} m_1 r_1 - m_2 r_2 &= 0 \\ m_1^2 r_1^2 + m_2^2 r_2^2 &= 2m_1 m_2 r_1 r_2. \end{aligned} \quad (1.4.183)$$

So clearly now

$$I = \frac{m_1 m_2}{m_1 + m_2} (r_1 + r_2)^2 \equiv \mu r^2. \quad (1.4.184)$$

L^2 is, from Equation 1.4.168,

$$L^2 = \hbar^2 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right]. \quad (1.4.185)$$

So we have our Hamiltonian:

$$H = -\frac{\hbar^2}{2I} \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right]. \quad (1.4.186)$$

The eigenproblem is just

$$\frac{1}{2I} L^2 \psi = E \psi. \quad (1.4.187)$$

We have already solved this, the eigenfunction is simply the spherical harmonic Y_J^m ⁶, and the allowed energies are

$$E_J = \frac{\hbar^2}{2I} J(J+1) \equiv B J(J+1), \quad J = 0, 1, 2, \dots \quad (1.4.188)$$

The degeneracy is the same $(2J+1)$ as the angular wavefunction.

todo: Selection rule and transition energies, priority: only when gone through p.177 of [13].

1.4.5 Spin

1.4.5.1 Formalism

Elementary particles carry **intrinsic** spins, *i.e.*, it is unrelated to actual rotation of matter. We also make the assertion that the operator formalism of spin is entirely analogous to

⁶We use J instead of ℓ as the bottom index here to avoid confusion with the azimuthal quantum number of the hydrogen atom, as is standard practice.

that of angular momentum, which is to say

Definition 1.4.5.1 (Commutation relations of spin operators).

$$[S_x, S_y] = i\hbar S_z \quad (1.4.189)$$

and cyclic permutations thereof.

And it follows that the *eigenvectors*⁷ of S^2 and S_z satisfy

Property 1.4.5.1 (Eigenfunctions of spin operators).

$$S^2|s, m\rangle = \hbar^2 s(s+1)|s, m\rangle, \quad S_z|s, m\rangle = \hbar m|s, m\rangle, \quad (1.4.190)$$

and

$$S_{\pm}|s, m\rangle = \hbar\sqrt{s(s+1) - m(m \pm 1)}|s, (m \pm 1)\rangle, \quad (1.4.191)$$

where

$$S_{\pm} \equiv S_x \pm iS_y. \quad (1.4.192)$$

We admit these values of s and m :

$$s = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots, \quad m = -s, (-s+1), \dots, (s-1), s. \quad (1.4.193)$$

s turns out to be an intrinsic and immutable value specific to each elementary particle: electrons have spin $1/2$, photons 1 , and so on.

1.4.5.2 Spin one half

We focus on the case of spin $1/2$, and the treatment can be easily generalised to other spins. Now there are just two eigenstates: **spin up** (\uparrow , $|1/2, 1/2\rangle$), and **spin down**, (\downarrow , $|1/2, -1/2\rangle$). At this point we assert that the general state of a spin $1/2$ particle can be expressed as a two-element column matrix called a **spinor**, whose elements can be complex: **todo-supo: can this even be derived?**

$$\chi = \begin{pmatrix} a \\ b \end{pmatrix} \equiv a\chi_+ + b\chi_-, \quad (1.4.194)$$

⁷We'll soon show that they are indeed vectors and not continuous functions like before

where

$$\chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (1.4.195)$$

representing spin up and down states.

Therefore spin operators *must* be 2×2 matrices, and we will work out their components now. We note that

$$S^2 \chi_+ = \frac{3}{4} \hbar^2 \chi_+, \quad S^2 \chi_- = \frac{3}{4} \hbar^2 \chi_-. \quad (1.4.196)$$

Therefore we need to solve the matrix equation

$$\underbrace{\begin{pmatrix} c & d \\ e & f \end{pmatrix}}_{S^2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{3}{4} \hbar^2 \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \quad (1.4.197)$$

Solving the equation gives

$$S^2 = \frac{3}{4} \hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (1.4.198)$$

Similar, we use

$$S_z \chi_+ = \frac{\hbar}{2} \chi_+, \quad S_z \chi_- = -\frac{\hbar}{2} \chi_- \quad (1.4.199)$$

to work out that

$$S_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1.4.200)$$

Similarly, from

$$= S_+ \chi_- = \hbar \chi_+, \quad S_- \chi_+ = \hbar \chi_-, \quad S_+ \chi_+ = S_- \chi_- = 0, \quad (1.4.201)$$

we get

$$S_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad S_- = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad (1.4.202)$$

And it immediately follows from definition that

$$S_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (1.4.203)$$

We can compactly write

Property 1.4.5.2 (Components of spin operators).

$$\mathbf{S} = \frac{\hbar}{2} \boldsymbol{\sigma}, \quad (1.4.204)$$

where

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1.4.205)$$

These σ 's are called **Pauli spin matrices**. And

$$S_x^2 = S_y^2 = S_z^2 = \frac{\hbar^2}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (1.4.206)$$

which is to say measurements of $S_{x_i}^2$ are sure to yield only $\hbar^2/4$.

The eigenspinors of S_z are χ_{pm} with eigenvalues $\pm\hbar/2$. So if we measure S_z on a particle in the general state (Equation 1.4.194), we might obtain either of $\pm\hbar/2$ with probabilities $|a|^2$ and $|b|^2$, therefore we must normalise:

$$|a|^2 + |b|^2 = 1, \quad (1.4.207)$$

where a and b can be complex.

If we want to know the outcomes of measuring S_x , we need to know its eigenvalues and eigenspinors. Solving the characteristic equation for S_x we have

$$\chi_+^{(x)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \chi_-^{(x)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad (1.4.208)$$

with eigenvalues $\pm\hbar/2$. This shouldn't be surprising as axes are arbitrarily oriented. Similarly, for S_y , the two eigenspinors are

$$\chi_{\pm}^{(y)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm i \end{pmatrix} \quad (1.4.209)$$

,

with the eigenvalues $\pm\hbar/2$ again of course.

As the eigenvectors of a hermitian matrix, they must span the space, that's why *any* generic

spinor χ can be expressed as a linear combination of these eigenspinors.

1.4.5.3 Addition of angular momenta - the intuitive way

Suppose we have two spin half particles, for example an electron and a proton as in the ground state of hydrogen atom ⁸. Each can take up spin up or down state so there are four possibilities in all:

$$\uparrow\uparrow, \uparrow\downarrow, \downarrow\uparrow, \downarrow\downarrow, \quad (1.4.210)$$

todo-supo: what's happening here? external product? where the first arrow refers to the electron and second the proton. We want to get the total angular momentum of the atom. Remembering that this is the ground state and $L = 0$, we let

$$\mathbf{S} \equiv \mathbf{S}^{(1)} + \mathbf{S}^{(2)}. \quad (1.4.211)$$

Each of these four composite states is an eigenstate of S_z as the z components simply add together:

$$\begin{aligned} S_z \chi_1 \chi_2 &= (S_z^{(1)} + S_z^{(2)}) \chi_1 \chi_2 \\ &= (S_z^{(1)} \chi_1) \chi_2 + \chi_1 (S_z^{(2)} \chi_2) \\ &= \hbar(m_1 + m_2) \chi_1 \chi_2. \end{aligned} \quad (1.4.212)$$

The quantum numbers for the composite system, $m = m_1 + m_2$, are

$$\begin{aligned} \uparrow\uparrow: m &= 1 \\ \uparrow\downarrow: m &= 0 \\ \downarrow\uparrow: m &= 0 \\ \downarrow\downarrow: m &= -1. \end{aligned} \quad (1.4.213)$$

Does this mean, then, that we have four states, with one transition $\Delta m = 0$? This is unphysical because if we apply the lowering operator to the top state we find:

$$\begin{aligned} S_- (\uparrow\uparrow) &= (S_-^{(1)} \uparrow) \uparrow + \uparrow (S_-^{(2)} \uparrow) \\ &= \hbar(\downarrow\uparrow + \uparrow\downarrow). \end{aligned} \quad (1.4.214)$$

Applying the lowering operator again we find

$$\begin{aligned} S_- (\downarrow\uparrow + \uparrow\downarrow) &\propto (S_-^{(1)} + S_-^{(2)}) (\downarrow\uparrow + \uparrow\downarrow) \\ &\propto \downarrow\downarrow. \end{aligned} \quad (1.4.215)$$

⁸This is so that the angular momentum is zero.

This means that there are only three states and the middle state is a superposition of two ‘mixed’ states, and as we see m runs from -1 to $+1$, which means that the total angular momentum $s = 1$

Property 1.4.5.3 (Triplet state). The three (triplet) states with $s = 1$ are

$$\begin{aligned} |1, 1\rangle &= \uparrow\uparrow \\ |1, 0\rangle &= \frac{1}{\sqrt{2}}(\uparrow\downarrow + \downarrow\uparrow) \text{ (symmetric wrt exchange of labels)} \\ |1, -1\rangle &= \downarrow\downarrow. \end{aligned} \quad (1.4.216)$$

However, we notice that for the state

$$|0, 0\rangle \equiv \frac{1}{\sqrt{2}}(\uparrow\downarrow - \downarrow\uparrow) \text{ (antisymmetric wrt exchange of labels),} \quad (1.4.217)$$

if we apply the raising or lowering operator we get zero. This means that this state has $s = 0$, therefore the notation.

Therefore we see that the combination of two spin half particles can carry a total spin of either 1 or 0. Therefore according to Equation 1.4.190, an eigenvalue of $2\hbar^2$ or 0 must be obtained if the triplet or singlet state is acted on by S^2 . We check if this is the case: first we write S^2 :

$$\begin{aligned} S^2 &= (\mathbf{S}^{(1)} + \mathbf{S}^{(2)}) \cdot (\mathbf{S}^{(1)} + \mathbf{S}^{(2)}) \\ &= (S^{(1)})^2 + (S^{(2)})^2 + 2\mathbf{S}^{(1)} \cdot \mathbf{S}^{(2)} \\ &= (S^{(1)})^2 + (S^{(2)})^2 + 2(S_x^{(1)}S_x^{(2)} + S_y^{(1)}S_y^{(2)} + S_z^{(1)}S_z^{(2)}). \end{aligned} \quad (1.4.218)$$

So

$$\begin{aligned} S^2|1, 1\rangle &= S^2 \uparrow\uparrow \\ &= \frac{3\hbar^2}{4}[\uparrow\uparrow + \uparrow\uparrow + 2(\downarrow\downarrow - \downarrow\downarrow + \uparrow\uparrow)] \\ &= 2\hbar^2, \end{aligned} \quad (1.4.219)$$

where we have used Equation 1.4.205 to get the effects of the operators on spinors.

Other triplet states and the singlet state can be checked similarly.

1.4.5.4 Addition of angular momenta - the ground-up way

The following section is adapted from Chapter 4 of [2]. We consider a system where there are two sources of angular momentum, j_1 and j_2 , which includes orbital and spin

momentum. The complete state of the system can be specified as

$$|j_1 m_{j1}; j_2 m_{j2}\rangle, \quad (1.4.220)$$

because the operators $J_1^2, J_2^2, J_{1z}, J_{2z}$ commute (m_j is the eigenvalue of J_z).

We now see that the total angular momentum, $j = j_1 + j_2$, can also be specified simultaneously. We first see that it is really an angular momentum by evaluating the commutators of its components (we know that the one particle j 's are angular momenta and we know their commutators):

$$\begin{aligned} [J_x, J_y] &= [J_{1x} + J_{2x}, J_{1y} + J_{2y}] \\ &= [J_{1x}, J_{1y}] + [J_{2x}, J_{2y}] + 0 + 0 \\ &= i\hbar J_{1z} + i\hbar J_{2z} \\ &= i\hbar J_z. \end{aligned} \quad (1.4.221)$$

This is characteristic of angular momentum. And so we know from the generalised (Theorem 1.4.3.2) properties of angular momentum that its allowed magnitude is⁹

$$\hbar\sqrt{j(j+1)}, \quad (1.4.222)$$

where j can take integer or half-integer values. Its z -component is

$$\hbar m_j, \quad m_j = -j, -j+1, \dots, 0, \dots, j-1, j. \quad (1.4.223)$$

Because J_1^2 and J_2^2 commutes with all their components and J^2 can be expressed in terms of those components, we can say

$$[J^2, J_i^2] = 0, \quad i = 1, 2. \quad (1.4.224)$$

So the total angular momentum can be specified simultaneously with individual angular momenta. Note that although we have found the *allowed* values of total angular momentum, we have yet to find how to calculate the actual value for a system with specified individual momenta.

⁹In case you get confused here, remember that j is but a quantum number, *i.e.*, an index, not the actual magnitude of the momentum in question.

We now want to know if we can specify m_{ji} and j together:

$$\begin{aligned}
 [J_{1z}, J^2] &= [J_{1z}, J_x^2] + [J_{1z}, J_y^2] + [J_{1z}, J_z^2] \\
 &= [J_{1z}, (J_{1x} + J_{2x})^2] + [J_{1z}, (J_{1y} + J_{2y})^2] + [J_{1z}, (J_{1z} + J_{2z})^2] \\
 &= [J_{1z}, J_{1x}^2 + 2J_{1x}J_{2x}] + [J_{1z}, J_{1y}^2 + 2J_{1y}J_{2y}] \\
 &= [J_{1z}, J_{1x}^2] + 2[J_{1z}, J_{1x}]J_{2x} + 2[J_{1z}, J_{1y}]J_{2y} \\
 &= [j_{1z}, J_1^2 - J_{1z}^2] + 2i\hbar J_{1y}J_{2x} - 2i\hbar J_{1x}J_{2y} \\
 &= 2i\hbar(J_{1y}J_{2x} - J_{1x}J_{2y}).
 \end{aligned} \tag{1.4.225}$$

So, no, we cannot specify m_{ji} if we specify j .

Therefore we have to choose between specifying the system as either

$$|j_1 m_{j1}; j_2 m_{j2}\rangle, \tag{1.4.226}$$

the **uncoupled picture**, or

$$|j_1 j_2; j m_j\rangle, \tag{1.4.227}$$

the **coupled picture**. For now the choice is equivalent and arbitrary but in future we will revisit this concept, such as in Section 3.3.2, where the coupled picture makes identification of triplet and singlet states easier.

1.4.5.5 Permitted values of total momentum

We now attempt to identify what the actual values of the total momentum quantum numbers j and m_j can be, amongst the range of allowed values.

Since $j_z = j_{1z} + j_{2z}$, we can immediately obtain

$$m_j = m_{j1} + m_{j2}, \tag{1.4.228}$$

that is, the z -components *simply add*.

The values of j is much thornier: the total number of states, given j_1 and j_2 , is

$$(2j_1 + 1)(2j_2 + 1) = 4j_1 j_2 + 2j_1 + 2j_2 + 1. \tag{1.4.229}$$

The maximum value of j is the maximum value of m_j , which is naturally $j_1 + j_2$. So there are

$$2j + 1 = 2j_1 + 2j_2 + 1 \tag{1.4.230}$$

degenerate states corresponding to the maximum total angular momentum. We have $4j_1 j_2$ more states to find. We descend m_j by 1 and note that there are two ways (actu-

ally linear combinations due to the indistinguishability) to form this state, $m_{j1} = j_1 - 1$ and $m_{j2} = j_2$ or the other way around. The state with $j = m_{j1} + m_{j2}$ already account for one of the linear combinations (as the degenerate state with m_j 1 less than the maximum), so there must be another *coupled state* with maximum $j = j_1 + j_2 - 1$, whose maximum m_j , of course, equal to the same value. This coupled state accounts for another $2j + 1 = 2j_1 + 2j_2 - 1$ states, and so on, and we can show that the sum **todo-supo: don't know how to proceed**

$$(2j_1 + 2j_2)N - 1 - 3 - 5 - \dots \quad (1.4.231)$$

It turns out by the time we have reached $j = |j_1 - j_2|$ all the states will be accounted for, and this is known as the **Clebsch-Gordan series** for allowed values of j :

$$j = j_1 + j_2, j_1 + j_2 - 1, \dots, |j_1 - j_2|. \quad (1.4.232)$$

The series can be very intuitively understood in Figure 1.1

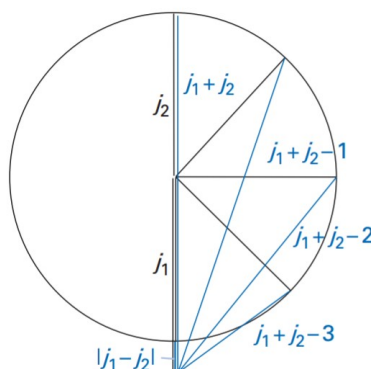


Figure 1.1: The allowed values of j are just the third side of the triangle.

1.4.5.6 Relation between coupling schemes

In Section 1.4.5.4 we have introduced two pictures of coupling, the uncoupled and the coupled picture. Just as a recap, the uncoupled picture specifies the individual quantum numbers j_i and m_{ji} . We usually use the **vector model of coupled angular momenta** to represent the differences between the two schemes, as in Figure 1.2.

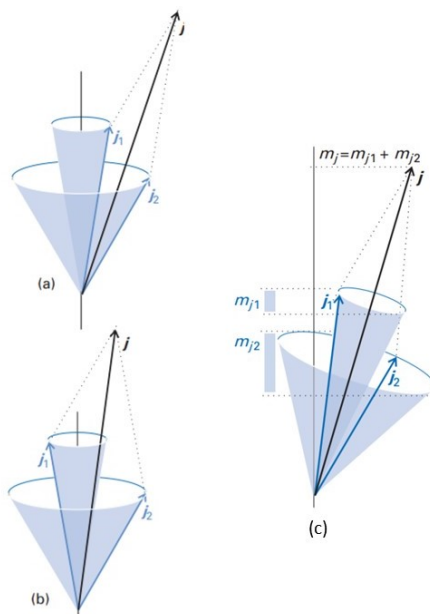


Figure 1.2: (a) and (b) represent the uncoupled picture, in which we can't specify j_i 's and j simultaneously but can do so with m_j , as the z -component is specified. (c) represents the coupled picture where the two j_i 's are locked together, so j (hence also m_j) is specified, but not the individual components that add up to m_j .

We are more interested in the coupled picture as it gives us more immediate information of the system total parameters, as we shall see in Section 3.3.2. So we should develop a way to quickly transform between the two pictures: the state $|j_1 j_2; j m_j\rangle$ is built from a linear combination of all states with $m_{j1} + m_{j2} = m_j$:

$$|j_1 j_2; j m_j\rangle = \sum_{m_{j1}, m_{j2}} C(m_{j1}, m_{j2}) |j_1 m_{j1}; j_2 m_{j2}\rangle, \quad (1.4.233)$$

where the coefficients $C(m_{j1}, m_{j2})$ are called **Clebsch-Gordan coefficients**, or vector coupling coefficients. For our case, $m_{j1} = m_{j2} = 1/2$, the coefficients are easily seen to agree with our earlier efforts at establishing the singlet and triplet states. Omitting $j_1 j_2$ in the notation and simply writing $|j, m_j\rangle$ or more specifically, $|S, M_S\rangle$, we have

$$\begin{aligned} |1, +1\rangle &= \alpha_1 \alpha_2 \\ |1, 0\rangle &= \frac{1}{\sqrt{2}} \alpha_1 \beta_2 + \frac{1}{\sqrt{2}} \beta_1 \alpha_2 \\ |1, -1\rangle &= \beta_1 \beta_2 \\ |0, 0\rangle &= \frac{1}{\sqrt{2}} \alpha_1 \beta_2 - \frac{1}{\sqrt{2}} \beta_1 \alpha_2. \end{aligned} \quad (1.4.234)$$

The Clebsch-Gordan coefficients can be calculated as overlap integrals between the coupled and uncoupled states. We do so by right-multiplying Equation 1.4.233 by $\langle j_1 m'_{j1}; j_2 m'_{j2}|$:

$$\langle j_1 m'_{j1}; j_2 m'_{j2} | j_1 j_2; j m_j \rangle = C(m'_{j1}, m'_{j2}), \quad (1.4.235)$$

where other terms in the sum vanishes by the orthogonality of the states. This can be intuitively understood as how much the coupled state resembles the uncoupled state, or 'how much' of the uncoupled state should be in the linear combination. For example, the state $|1, +1\rangle$ must be composed of $\alpha_1 \alpha_2$ only as this is the only state with $M_S = +1$.

1.4.5.7 Atomic term symbols

read chap 7 of atkins, and section on spin in B field in griffiths, priority: high. After that, re-write this section

Term symbols give better descriptions than electron configurations, as we will see, same electronic configuration will give rise to different energies as a result of electrostatic interactions. The idea is to determine the total orbital angular momentum L and the total spin angular momentum S and then adding them vectorially to get the total angular momentum J :

$$\mathbf{J} = \mathbf{L} + \mathbf{S} = \sum_i \mathbf{l}_i + \sum_j \mathbf{s}_j. \quad (1.4.236)$$

This is called the **Russell-Saunders coupling** or **LS coupling**, which is only predominant in lighter ($Z < 30$) atoms. same as spin-orbit coupling? The result of this sum is summarised as a **term symbol**, like

$$^{2S+1}L_J, \quad (1.4.237)$$

where L, S, J are the total orbital/total spin/total angular momentum quantum numbers, and $2S + 1$ is known as the spin multiplicity.

With an argument exactly like the one given in Section 1.4.5.5, we can show that \mathbf{J} and \mathbf{S} are angular momenta and their values must run through integers and/or half integers. And since individual orbital angular momenta must be integers and individual spins can be both, so do the total. And another argument as in Section 1.4.5.5 will give us that J runs from $L + S$ to $|L - S|$. To summarise, the allowed values for the quantum numbers involved in the term symbol are

Property 1.4.5.4 (Term symbol). The term symbol is

$$^{2S+1}L_J, \quad (1.4.238)$$

where the allowed values are

$$\begin{aligned} S &= 0, 1/2, 1, 3/2, \dots \\ (2S + 1) &= 1, 2, 3, 4, \dots \\ L &= 0, 1, 2, 3, \dots \\ J &= L + S, L + S - 1, \dots, |L - S|. \end{aligned} \quad (1.4.239)$$

In analogy to assigning s, p, d, f, \dots to $\ell = 0, 1, 2, 3, \dots$, we assign S, P, D, F, \dots for $L = 0, 1, 2, 3, \dots$.

Recall that the z -components of angular momenta simply add, *i.e.*,

$$\begin{aligned} L_z &= \sum_i l_{zi} = \sum_i m_{li} = M_L \\ S_z &= \sum_i s_{zi} = \sum_i m_{si} = M_S \end{aligned} \quad (1.4.240)$$

The M_S quantum number gives rise the $2S + 1$ multiplicity (their energy will only be further split into levels in presence of external field **todo: read up relevant stuff, reference, priority: low**). We will now look at a few examples.

Example 1: ns^2

- s orbital means $\ell = 0$, so $m_{li} = 0$, and $M_L = 0$, and this being the only allowed value of M_L means $L = 0$ or the S-state.
- The two spins must be antiparallel, so M_S is also equal 0, and so similarly $S = 0$.
- By the same token, J can also only be 0.
- Therefore, the term symbol for ns^2 is 1S_0 .

Example 2: np^6

- p orbital means $\ell = 1$, so $m_{li} = -1, 0, 1$ so M_L being the sum of all occupied orbitals (which is all of them) is 0.
- Spins also all pair up to give $M_S = 0$.

- Therefore J can only be 0
- The term symbol for np^6 , and actually for **all** fully filled orbitals subshells, is 1S_0 . It's also called **singlet S zero**.

Example 3: $ns^1n's^1$ (first excited state of Helium)

- M_L can only be zero as $m_{li} = 0$.
- However now the spins are not confined to have antiparallel spins, and can independently take on values of $\pm 1/2$, so M_S can be $-1, 0, 1$. This means the largest value of S is 1 since that's the largest value M_S can take on. So we must have $S = 0, 1$, corresponding to 3S (triplet) and 1S (singlet) states.

We list the possible states below:

| M_L | M_S | | |
|-------|------------|----------------------|------------|
| | 1 | 0 | -1 |
| 0 | $0^+, 0^+$ | $0^+, 0^-; 0^-, 0^+$ | $0^-, 0^-$ |

where 0^+ means $m_l = 0$ and $m_s = +1/2$ and vice versa. The middle two states are not indistinguishable because the spatial orbitals are not the same ($1s$ and $2s$) There are 4 **microstates**. The triplet state accounts for one state from each column, and the singlet state must claim the remaining M_S state, and it doesn't matter which one each state takes.

And we are left with J to determine. We add M_L and M_S to get M_J . For the triplet state $M_J, M_J = 1, 0, -1$, so $J = 1$. For the singlet state similarly $J = 0$.

Therefore the two term symbols the configuration $ns^1n's^1$ correspond to are 3S_1 and 1S_0 .

Example 4: Carbon atom

As shown previously, we do not need to worry about fully filled subshells and only need to focus on $2p^2$. Let's derive a general result first:

Lemma 1.4.5.1 (Number of electron assignments). For G number of **equivalent** (same subshell) spin-orbitals and N electrons, which are indistinguishable, the number of ways to assign the electrons, $D(G, N)$ is

$$D(G, N) = \frac{G!}{N!(G-N)!}. \quad (1.4.241)$$

So for carbon, we have 15 ways to assign the electrons. Intuitively, the first electron can take any of the 6 spin-orbitals, the second can take the remaining 5, and a factor of $2!$ is divided through for they are indistinguishable, giving a total of 15 ways to arrange.

To find all these 15 microstates we again use a table. But before that we need to find out the values of M_L and M_S . It's easily seen that M_L runs from -2 to $+2$ and M_S from -1 to $+1$. So $L = 0, 1, 2$ and $S = 0, 1$. The table is then

| M_L | M_S | | |
|-------|-----------------------|----------------------------------|-----------------------|
| | 1 | 0 | -1 |
| 2 | $1^+, 1^+$ | $1^+, 1^-$ | $1^-, 1^-$ |
| 1 | $0^+, 1^+$ | $1^+, 0^-; 1^-, 0^+$ | $0^-, 1^-$ |
| 0 | $0^+, 0^+; 1^+, -1^+$ | $1^+, -1^-; -1^+, 1^-; 0^+, 0^-$ | $1^-, -1^-; 0^-, 0^-$ |
| -1 | $0^+, -1^+$ | $0^+, -1^-; 0^-, -1^+$ | $0^-, -1^-$ |
| -2 | $-1^+, -1^+$ | $-1^+, -1^-$ | $-1^-, -1^-$ |

where we make no distinction like the previous case between $1^+, 0^-$ and $0^-, 1^+$ because the spatial orbitals are equivalent now. The grey terms are forbidden by the Pauli exclusion principle. We now assign term symbols to the microstates.

We start by remarking that although the **allowed** values of L and S is 0 to 2 and 0 to 1 respectively, the actual **permitted** term symbols are not necessarily all possible permutations of these values. For example, the largest value of M_L is 2 and it only occurs with $M_S = 0$, this means that $L = 2$ state only occurs with $S = 0$. For the remaining states, the largest $M_L = 1$, and it occurs with all three M_S values, hence there must be a state with $L = 1$ and $S = 1$. We list the term symbols below:

- $L = 2, S = 0$, i.e., 1D : one state per row in the $M_S = 0$ column. J can only take up one value, which is 2. The complete term symbol is 1D_2 .
- $L = 1, S = 1$, i.e., 1P : nine states per remaining cells. J can take up values from 2 to 0, so complete term symbols are $^3P_2, ^3P_1$, and 3P_0 .¹⁰
- $L = 0, S = 0$, term symbol 1S_0 , naturally.

As can be seen, we don't really have to specify J as it can be deduced from S and L . The degeneracies of each state is given by $2J + 1$.

At last, we can introduce Hund's rules, **todo: theoretical foundations, priority: medium:**

¹⁰Another way to look at this is that the nine states correspond to M_J of 2, 1, 1, 0, 0, -1, 0, -1, -2, so we can isolate three sets of M_J running from $-J$ to J with $J = 0, 1, 2$ respectively.

Theorem 1.4.5.1 (Hund's rules). The three Hund's rules state that

1. The state with the largest value of S is the most stable (has the lowest energy), and stability decreases with decreasing S .
2. For states with the same value of S , the state with the largest value of L is the most stable.
3. If the states have the same value of L and S , then, for a subshell that is less than half filled, the state with the smallest value of J is the most stable; for a subshell that is more than half filled, the state with the largest value of J is the most stable.

1.4.5.8 Molecular term symbols

The following discussion only strictly applies to homonuclear diatomics, for which only the orbital angular momentum along the internuclear axis is defined.

The molecular term symbol is given in the form of

$$^{2\Sigma+1}\Lambda_{g/u}^{\pm}, \quad (1.4.242)$$

Σ is just the spin angular momentum, effectively the half the number of unpaired electrons in the molecule.

Λ is the orbital angular momentum, given as

$$\Lambda = \sum_i M_{L,i}, \quad (1.4.243)$$

with *unpaired* electrons in σ orbitals having $M_L = 0$ and those in π orbitals having $M_L = 1$ and so on.

The g/u labels denotes the *inversion symmetry*, where the overall inversion symmetry is a direct product of the inversion symmetry of all occupied orbitals, with

$$g \times g = g \quad u \times u = g \quad g \times u = u \quad u \times g = u. \quad (1.4.244)$$

And finally the \pm label denotes the *mirror-plane* symmetry, where the mirror plane contains the internuclear axis, where $+$ means unchanged upon reflection and $-$ means the sign changes upon reflection. Mathematically speaking, reflection changes the sign of ϕ . Let's

see this in action, say we have the ground state of oxygen, which has an electronic configuration of π^2 , which has two orbitals with $M_L = \pm 1$, which leads to two $^1\Delta$ states with two electrons in either orbitals, and a $^3\Sigma$ and $^1\Sigma$ state with two electrons in both states but two orientations. The wavefunctions for the two orbitals are

$$\begin{aligned}\pi_+ &= F(\mathbf{r}_a, \mathbf{r}_b)e^{i\phi} \\ \pi_- &= F(\mathbf{r}_a, \mathbf{r}_b)e^{-i\phi}\end{aligned}\tag{1.4.245}$$

we further define the reflection operator

$$\sigma_{xy}\pi_{\pm} = \pi_{\mp}\tag{1.4.246}$$

and the total wavefunction for the two electrons

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}}[\pi_+(1)\pi_-(2) \mp \pi_-(1)\pi_+(2)] \times \sigma_{\pm}\tag{1.4.247}$$

where the top sign denotes triplet, whose spin wavefunction is symmetric to exchange, and the bottom sign denotes singlet. Now the action of the total reflection operator $\sigma_{\text{tot}} = \sigma_{xy}(1)\sigma_{xy}(2)$ is

$$\begin{aligned}\sigma_{\text{tot}}\Psi(\mathbf{r}_1, \mathbf{r}_2) &= \frac{1}{\sqrt{2}}\sigma_{xy}(1)\sigma_{xy}(2)[\pi_+(1)\pi_-(2) \mp \pi_-(1)\pi_+(2)] \times \sigma_{\pm} \\ &= \frac{1}{\sqrt{2}}[\pi_-(1)\pi_+(2) \mp \pi_+(1)\pi_-(2)] \times \sigma_{\pm} \\ &= \mp\Psi(\mathbf{r}_1, \mathbf{r}_2)\end{aligned}\tag{1.4.248}$$

therefore we see that the final full term symbols are $^3\Sigma^-$ and $^1\Sigma^+$.

This is only going to matter in $\Lambda = 0$ (Σ) states, this is because all other states have a two-fold degeneracy, so the reflection operation is just going to interchange them, whereas in Σ states there's no degeneracy.

More examples

$1\sigma^1 2\sigma^1$

As the electrons don't occupy the same orbitals, Pauli exclusion principle doesn't apply anyway, and also, since σ MOs have no ϕ dependence, it has to be unchanged upon reflection, therefore $^1\Sigma^+$ and $^3\Sigma^+$ are produced.

σ^1

Similar argument produces $^2\Sigma^+$.

Closed shell

Close shells must have $S = 0$ which forces the singlet configuration, which means anti-symmetric spin wavefunction and hence $^1\Sigma^+$.

π^1

This simply produces $^2\Pi$, specifying the signs is unnecessary as the term is 2-degenerate anyway.

1.5 IDENTICAL PARTICLES

1.5.1 Two particle systems

1.5.1.1 The Schrödinger equation

The state of a two-particle system is a function of both particle's coordinates and time:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, t). \quad (1.5.1)$$

Its time evolution is, as always, determined by the Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial t} = H\Psi, \quad (1.5.2)$$

where the Hamiltonian is

$$H = -\frac{\hbar^2}{2m_1} \nabla_1^2 - \frac{\hbar^2}{2m_2} \nabla_2^2 + V(\mathbf{r}_1, \mathbf{r}_2, t) \quad (1.5.3)$$

The probability of finding particle 1 in $d^3\mathbf{r}_1$ and particle 2 in $d^3\mathbf{r}_2$ is

$$|\Psi(\mathbf{r}_1, \mathbf{r}_2, t)|^2 d^3\mathbf{r}_1 d^3\mathbf{r}_2. \quad (1.5.4)$$

The normalisation requirement is therefore

$$\int |\Psi(\mathbf{r}_1, \mathbf{r}_2, t)|^2 d^3\mathbf{r}_1 d^3\mathbf{r}_2 = 1. \quad (1.5.5)$$

For **time-independent** potentials, we can use the usual separation of variables:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, t) = \psi(\mathbf{r}_1, \mathbf{r}_2) e^{-iEt/\hbar}, \quad (1.5.6)$$

where ψ is such that

$$H\psi = E\psi. \quad (1.5.7)$$

1.5.1.2 Centre-of-mass coordinates

If, in a **two particle system**, where the potential depends only on $\mathbf{r} \equiv \mathbf{r}_1 - \mathbf{r}_2$ between the two particles, we can use the centre-of-mass coordinates to simplify and *separate* the

Schrödinger equation as follows: we define the **centre-of-mass coordinate**:

$$\mathbf{R} \equiv \frac{m_1 \mathbf{r}_1 + m_2 \mathbf{r}_2}{m_1 + m_2} \quad (1.5.8)$$

Now, we need to re-write the Schrödinger equation in terms of the new coordinates, and first of all we need to write out the Hamiltonians, which is fiddly to do: we first need to see that

$$\nabla_1 f = \frac{\partial f}{\partial r_{1x}} + \frac{\partial f}{\partial r_{1y}} + \frac{\partial f}{\partial r_{1z}}. \quad (1.5.9)$$

So to re-write it, we note that r_{1x} is only dependent on r_x and R_x , and so on, so

$$\begin{aligned} \frac{\partial f}{\partial r_{1x}} &= \frac{\partial f}{\partial r_x} \frac{\partial r_x}{\partial r_{1x}} + \frac{\partial f}{\partial R_x} \frac{\partial R_x}{\partial r_{1x}} \\ &= \frac{\partial f}{\partial r_x} + \frac{\partial f}{\partial R_x} \frac{\mu}{m_2}, \end{aligned} \quad (1.5.10)$$

and

$$\begin{aligned} \frac{\partial f}{\partial r_{2x}} &= \frac{\partial f}{\partial r_x} \frac{\partial r_x}{\partial r_{2x}} + \frac{\partial f}{\partial R_x} \frac{\partial R_x}{\partial r_{2x}} \\ &= -\frac{\partial f}{\partial r_x} + \frac{\partial f}{\partial R_x} \frac{\mu}{m_1}, \end{aligned} \quad (1.5.11)$$

and so on for other components, so

$$\nabla_1 = \nabla_r + \frac{\mu}{m_2} \nabla_R \quad (1.5.12a)$$

$$\nabla_2 = -\nabla_r + \frac{\mu}{m_1} \nabla_R. \quad (1.5.12b)$$

To get the Laplacian, we note that it's not equivalent to left multiplying the transformed nabla operator twice, but instead we have to go back to Equation 1.5.10, and differentiating it again wrt r_{1x} and so on:

$$\begin{aligned} \frac{\partial^2 f}{\partial (r_{1x})^2} &= \left(-\frac{\partial}{\partial r_x} + \frac{\mu}{m_2} \frac{\partial}{\partial R_x} \right) \left(-\frac{\partial f}{\partial r_x} + \frac{\partial f}{\partial R_x} \frac{\mu}{m_2} \right) \\ &= \frac{\partial^2 f}{\partial (r_x)^2} - \frac{\partial^2 f}{\partial r_x \partial R_x} \frac{2\mu}{m_2} + \frac{\partial^2 f}{\partial (R_x)^2} \left(\frac{\mu}{m_2} \right)^2, \end{aligned} \quad (1.5.13)$$

and

$$\begin{aligned} \frac{\partial^2 f}{\partial (r_{2x})^2} &= \left(\frac{\partial}{\partial r_x} + \frac{\mu}{m_1} \frac{\partial}{\partial R_x} \right) \left(\frac{\partial f}{\partial r_x} + \frac{\partial f}{\partial R_x} \frac{\mu}{m_1} \right) \\ &= \frac{\partial^2 f}{\partial (r_x)^2} - \frac{\partial^2 f}{\partial r_x \partial R_x} \frac{2\mu}{m_1} + \frac{\partial^2 f}{\partial (R_x)^2} \left(\frac{\mu}{m_1} \right)^2, \end{aligned} \quad (1.5.14)$$

so we have

$$-\frac{\hbar^2}{2m_1} \nabla_1^2 = -\frac{\hbar^2}{2m_1} \nabla_r^2 - \nabla_R^2 \frac{\hbar^2 \mu^2}{m_1 m_2^2} - (\text{mixed partial term}) \quad (1.5.15)$$

and

$$-\frac{\hbar^2}{2m_1}\nabla_2^2 = -\frac{\hbar^2}{2m_2}\nabla_r^2 - \nabla_R^2 \frac{\hbar^2\mu^2}{m_2m_1^2} + (\text{same mixed partial term}). \quad (1.5.16)$$

Finally, the Schrödinger equation is

$$\left[-\frac{\hbar^2}{2(m_1 + m_2)}\nabla_R^2 - \frac{\hbar^2}{2\mu}\nabla_r^2 + V(\mathbf{r}) \right] \psi = E\psi. \quad (1.5.17)$$

We attempt separation of variables:

$$\psi(\mathbf{R}, \mathbf{r}) = \psi_R(\mathbf{R})\psi_r(\mathbf{r}), \quad (1.5.18)$$

and we obtain two equations:

$$-\frac{\hbar^2}{2(m_1 + m_2)}\nabla_R^2\psi_R = E_R\psi_R \quad (1.5.19a)$$

$$\left[-\frac{\hbar^2}{2\mu}\nabla_r^2 + V(\mathbf{r}) \right] \psi_r = E_r\psi_r. \quad (1.5.19b)$$

This is saying that we have successfully decoupled the equation into two independent systems: the total mass $(m_1 + m_2)$, moving as a free particle (0 potential) and the relative motion as a single particle with reduced mass subject to potential V .

1.5.1.3 Bosons and fermions

Suppose we now have a two particle system that is non-interacting, which is to say that the two particles are in their own one-particle states and that they are independent - this is by no means always true, as in the singlet spin configuration where two spins are correlated - we can write

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2). \quad (1.5.20)$$

In writing this we have presumed that the two particles are **distinguishable**, otherwise the labels a and b wouldn't make any sense. However, almost all the particles we will be dealing with will *not* be distinguishable. Therefore, the most we can say is that *one* particle is in state a and the other in state b . We note that since they are indistinguishable, the exchange of labels *must* have no effect on the probability density, *i.e.*,

$$|\psi(1, 2)|^2 = |\psi(2, 1)|^2, \Rightarrow \psi(1, 2) = \pm\psi(2, 1). \quad (1.5.21)$$

So this means we can write

$$\psi_{\pm}(\mathbf{r}_1, \mathbf{r}_2) = A[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) \pm \psi_b(\mathbf{r}_1)\psi_a(\mathbf{r}_2)]. \quad (1.5.22)$$

This wavefunction is *non-committal* as to which particle is in which state, but when do we use which sign? We introduce the postulate

Postulate 1.5.1.1 (Bosons and fermions).

All particles with *integer* spins are bosons, and
all particles with *half integer* spins are fermions.

We use the *plus sign* for bosons and *minus sign* for fermions.

It then follows that

Theorem 1.5.1.1 (Pauli exclusion principle). Two identical fermions, for example, two electrons, cannot occupy the same state, for if $\psi_a = \psi_b$, then

$$\psi_{-}(\mathbf{r}_1, \mathbf{r}_2) = A[\psi_a(\mathbf{r}_1)\psi_a(\mathbf{r}_2) - \psi_a(\mathbf{r}_1)\psi_a(\mathbf{r}_2)] = 0, \quad (1.5.23)$$

and we are left with no wavefunction at all.

A more general re-formulation of problem requires the introduction of the **exchange operator**:

Definition 1.5.1.1 (Exchange operator). The exchange operators is defined as

$$Pf(\mathbf{r}_1, \mathbf{r}_2) = f(\mathbf{r}_2, \mathbf{r}_1). \quad (1.5.24)$$

Definition 1.5.1.2 (Properties of the exchange operator).

(1) It is immediately clear that

$$P^2 = I. \quad (1.5.25)$$

(2) And therefore the eigenvalues of P are ± 1 .

(3) If we have two *identical* particles, the Hamiltonian must treat them the same way: $m_1 = m_2$ and $V(\mathbf{r}_1, \mathbf{r}_2) = V(\mathbf{r}_2, \mathbf{r}_1)$, therefore P and H are compatible observables:

$$[P, H] = 0. \quad (1.5.26)$$

Theorem 1.5.1.2 (Symmetrisation requirement). Because P and H commute, we can find simultaneous eigenstates of both, *i.e.*, we can find solutions to Schrödinger equation that are either symmetric (eigenvalue $+1$) or antisymmetric (eigenvalue -1) under exchange of labels:

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \pm \psi(\mathbf{r}_2, \mathbf{r}_1). \quad (1.5.27)$$

Therefore, if a system begins in a state it must remain in this state. The plus sign is for bosons and negative for fermions.

1.5.1.4 Exchange forces

We examine a 1D example. Let's suppose one particle is in $\psi_a(x)$ and another in $\psi_b(x)$, and these two states are orthogonal and normalised. If the two particles are distinguishable, and number 1 is the one in state ψ_a , then the composite wavefunction is

$$\psi(x_1, x_2) = \psi_a(x_1)\psi_b(x_2). \quad (1.5.28)$$

If they are identical bosons or fermions, the composite will be

$$\psi_{\pm}(x_1, x_2) = \frac{1}{\sqrt{2}}[\psi_a(x_1)\psi_b(x_2) \pm \psi_b(x_1)\psi_a(x_2)]. \quad (1.5.29)$$

Let's calculate the expectation value of the square of the separation distance between the two particles, given by

$$\langle (x_1 - x_2)^2 \rangle = \langle x_1^2 \rangle + \langle x_2^2 \rangle - 2\langle x_1 x_2 \rangle. \quad (1.5.30)$$

Case 1: Distinguishable particles

We use the wavefunction in Equation 1.5.28:

$$\langle x_1 \rangle = \int x_1^2 |\psi_a(x_1)|^2 dx_1 \int |\psi_b(x_2)|^2 dx_2 = \langle x^2 \rangle_a, \quad (1.5.31)$$

(the expectation of x^2 in the *one-particle* state ψ_a) so similarly

$$\langle x_2^2 \rangle = \langle x_2^2 \rangle_b. \quad (1.5.32)$$

And

$$\langle x_1 x_2 \rangle = \int x_1 |\psi_a(x_1)|^2 dx_1 \int x_2 |\psi_b(x_2)|^2 dx_2 = \langle x \rangle_a \langle x \rangle_b. \quad (1.5.33)$$

Collecting, we have

$$\langle (x_1 - x_2)^2 \rangle = \langle x^2 \rangle_a + \langle x^2 \rangle_b - 2\langle x \rangle_a \langle x \rangle_b. \quad (1.5.34)$$

Case 2: Identical particles

We now have to use the wavefunctions in Equation 1.5.29:

$$\begin{aligned} \langle x_1^2 \rangle &= \frac{1}{2} \left[\int x_1^2 |\psi_a(x_1)|^2 dx_1 \int |\psi_b(x_2)|^2 dx_2 \right. \\ &\quad + \int x_1^2 |\psi_b(x_1)|^2 dx_1 \int |\psi_a(x_2)|^2 dx_2 \\ &\quad \pm \int x_1^2 \psi_a(x_1)^* \psi_b(x_1) dx_1 \int \psi_b(x_2)^* \psi_a(x_2) dx_2 \\ &\quad \left. \pm \int x_1^2 \psi_b(x_1)^* \psi_a(x_1) dx_1 \int \psi_a(x_2)^* \psi_b(x_2) dx_2 \right] \\ &= \frac{1}{2} [\langle x^2 \rangle_a + \langle x^2 \rangle_b \pm 0 \pm 0] \\ &= \frac{1}{2} (\langle x^2 \rangle_a + \langle x^2 \rangle_b). \end{aligned} \quad (1.5.35)$$

And likewise

$$\langle x_2^2 \rangle = \frac{1}{2} (\langle x^2 \rangle_b + \langle x^2 \rangle_a). \quad (1.5.36)$$

We observe that

$$\langle x_1^2 \rangle = \langle x_2^2 \rangle \quad (1.5.37)$$

since we can't tell them apart. Now

$$\begin{aligned} \langle x_1 x_2 \rangle &= \frac{1}{2} \left[\int x_1 |\psi_a(x_1)|^2 dx_1 \int x_2 |\psi_b(x_2)|^2 dx_2 \right. \\ &\quad + \int x_1 |\psi_b(x_1)|^2 dx_1 \int x_2 |\psi_a(x_2)|^2 dx_2 \\ &\quad \pm \int x_1 \psi_a(x_1)^* \psi_b(x_1) dx_1 \int x_2 \psi_b(x_2)^* \psi_a(x_2) dx_2 \\ &\quad \left. \pm \int x_1 \psi_b(x_1)^* \psi_a(x_1) dx_1 \int x_2 \psi_a(x_2)^* \psi_b(x_2) dx_2 \right] \\ &= \frac{1}{2} (\langle x \rangle_a \langle x \rangle_b + \langle x \rangle_b \langle x \rangle_a \pm \langle x \rangle_{ab} \langle x \rangle_{ba} \pm \langle x \rangle_{ba} \langle x \rangle_{ab}) \\ &= \langle x \rangle_a \langle x \rangle_b \pm |\langle x \rangle_{ab}|^2, \end{aligned} \quad (1.5.38)$$

where

$$\langle x \rangle_{ab} \equiv \int x \psi_a(x)^* \psi_b(x) dx. \quad (1.5.39)$$

Collecting we have

$$\langle (x_1 - x_2)^2 \rangle_{\pm} = \langle x^2 \rangle_a + \langle x^2 \rangle_b - 2\langle x \rangle_a \langle x \rangle_b \mp 2|\langle x \rangle_{ab}|^2. \quad (1.5.40)$$

Theorem 1.5.1.3 (Exchange force).

By comparing Equation 1.5.34 and Equation 1.5.40, we see that the difference lies in the final term:

$$\langle(\Delta x)^2\rangle_{\pm} = \langle(\Delta x)^2\rangle_d \mp 2|\langle x\rangle_{ab}|^2, \quad (1.5.41)$$

where the subscript d stands for ‘distinguishable’. Therefore we can see that identical bosons are closer together and identical fermions further apart than distinguishable particles in the same two states. Notice that $\langle x\rangle_{ab}$ **vanishes** unless the two wavefunctions **overlap**. Therefore, it is acceptable to treat electrons with non-overlapping wavefunctions as distinguishable. The increase or decrease in between identical particles are called an **exchange force**.

We now take a look at the hydrogen molecule, H_2 . To a good approximation¹¹, the ground state consists of one electron in the hydrogen atom ground state centred around nucleus 1 and another around nucleus 2. If electrons were bosons, the exchange force would concentrate the electrons towards the middle of the internuclear space, and as a result pull the protons inward, accounting for the covalent bond. However, electrons are fermions, and that would mean the concentration of negative charge should be on the wings, tearing the molecule apart. This is obviously wrong, and we are reminded that we have been ignoring spins all this while - the symmetrisation requirement (Theorem 1.5.1.2) requires that the **complete wavefunctions** of fermions be antisymmetric, not just the spatial wavefunctions. The complete wavefunctions of electrons are

$$\psi_e(\mathbf{r}, \mathbf{s}) = \psi(\mathbf{r})\chi(\mathbf{s}). \quad (1.5.42)$$

This is assuming that the spin and the spatial wavefunction are *uncoupled*, therefore separable. Under the same assumption, we can write the two-electron state as

$$\psi_{2e}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{s}_1, \mathbf{s}_2) = \psi_{\pm}(\mathbf{r}_1, \mathbf{r}_2)\chi(\mathbf{s}_1, \mathbf{s}_2), \quad (1.5.43)$$

where $\chi(\mathbf{s}_1, \mathbf{s}_2)$ can be one of the four (three triplet one singlet) states in Section 1.4.5.3. Now, we must require bonding, therefore symmetric¹², *spatial* wavefunctions. We remember that the singlet state is antisymmetric¹³, therefore the spatial wavefunction of the electron that’s multiplied to it can (and must) be symmetric. This can be confusing because

¹¹See next section.

¹²Remember that the exchange force favours bonding as long as the *spatial* wavefunction is symmetric, and it makes no use of the spinor as the ‘integration’ of orthonormal spinors always return the Kronecker delta.

¹³We can actually derive this from symmetry requirements alone, stating from the four possible combinations and then consider their coefficients as in [11].

we appear to have just derived that all electron wavefunctions must be antisymmetric. But in fact we have not, Theorem 1.5.1.2 only requires that the complete wavefunction be accordingly symmetrised, not spatial or spinor parts separately.

Therefore, we have now shown that the bonding orbital will have electrons occupy the singlet configuration, with total spin zero, and the antibonding orbital one of the triplet configurations.

Another more elementary but more intuitive way to look at this is that electrons with the same spin cannot be found at the same spot and the electron density will be low around another electron with the same spin. This is not the case for electrons with opposite spin. Because of the Coulombic repulsion, electrons with like spins are lower in energy than those with opposite spin. This explains in part the Hund's rule. Also, the triplet state is more stable than the singlet state, however the latter is essential for chemical bonding to happen, and bonding happens because the lowering in energy more than compensates for the exchange force.

Chapter 2

Spectroscopy

2.1 ATOMIC SPECTROSCOPY

2.1.1 Selection rules

Theorem 2.1.1.1 (The Laporte selection rule). The Laporte selection rule states that electric dipole transition that does not involve a change in parity is strictly forbidden, while those involving a change in parity may be allowed. In other words, the only transitions allowed are those involving a change in parity.

Proof. The transition probability between two states are given by the Einstein ‘B’ coefficient which is given in eq. 2.2.29, which is proportional to the modulus squared of the *transition dipole moment*

$$\mu_{i \rightarrow f} = \langle i | \hat{\mu} | f \rangle \quad (2.1.1)$$

Theorem 4.3.1.1 tells that the integrand must contain the tot. sym. irrep of the group the particle belongs to, which in this case is the full rotation group R_3 . The inversion operation¹ is diagnostic: if the integrand is antisymmetrical wrt. inversion the integral is necessarily zero. **Parity** is just the g/u label, but represented numerically as ± 1 . The parity of atomic orbitals is $(-1)^l$, because the dipole moment operator has u symmetry, the overall parity of the integrand is

$$(-1)^{l_i}(-1)(-1)^{l_f} \quad (2.1.2)$$

¹Just because we can readily deduce the parity from the orbital angular momentum. In principle any other operation is fine, we just need to see whether the direct product is $[1, 1, \dots, 1]$

which shows the two orbitals must have *opposite parity* for the integral to be plausibly non-zero. \square

This allows us to deduce the selection rule for orbital angular momentum of an atom

Theorem 2.1.1.2 (Selection rule for orbital angular momentum). The selection rule is

$$\Delta l = \pm 1 \quad (2.1.3)$$

Proof. As the dipole moment operator transforms like a translation which is exactly like $l = 1$ spherical harmonics, the representative of $\hat{\mu}$ is $\Gamma^{(1)}$. The overall direct product for $\langle i|\hat{\mu}|f\rangle$ is

$$\begin{aligned} & \Gamma^{(l_i)} \otimes \Gamma^{(1)} \otimes \Gamma^{(l_f)} \\ &= \left(\Gamma^{(l_i+1)} \oplus \Gamma^{(l_i)} \oplus \Gamma^{(l_i-1)} \right) \otimes \Gamma^{(l_f)} \end{aligned} \quad (2.1.4)$$

The requirement is for the integrand to *contain* the tot. sym. irrep, so clearly $\Gamma^{(l_f)}$ must be one of the three terms in front. And $\Delta l = 0$ is Laporte forbidden so we are left with the selection rule of $\Delta l = \pm 1$. \square

The selection rule for m_l can be deduced as well:

Theorem 2.1.1.3 (Selection rule for magnetic quantum number). The selection rule is

$$\Delta m_l = 0, \pm 1 \quad (2.1.5)$$

Proof. We now just need to look at the ϕ dependence only, which gives us the integral of the form

$$\int_0^{2\pi} e^{-im_{li}\phi} (\hat{\mu}_j) e^{im_{lf}\phi} d\phi \quad (2.1.6)$$

with $\hat{\mu} = -e\mathbf{r}$ is the electric dipole operator, which in spherical coordinates is

$$-e\mathbf{r} \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{pmatrix} \quad (2.1.7)$$

Now considering radiation that is plane-polarised with the electric field in the z -direction, remembering the first bra represents the complex conjugate, we get

$$\int_0^{2\pi} e^{-im_{lf}\phi} (-e\mathbf{r} \cos \theta) e^{im_{li}\phi} d\phi \propto \int_0^{2\pi} e^{i(m_{li}-m_{lf})\phi} d\phi \quad (2.1.8)$$

which gives $\Delta m_l = 0$. If the radiation is polarised in the x -direction, we have

$$\begin{aligned} & \int_0^{2\pi} e^{-im_{li}\phi} (-er \sin \theta \cos \phi) e^{im_{lf}\phi} d\phi \\ & \propto \int_0^{2\pi} e^{-im_{li}\phi} (e^{i\phi} + e^{-i\phi}) e^{im_{lf}\phi} d\phi \\ & = \int_0^{2\pi} e^{i(m_{lf}-m_{li}+1)\phi} + e^{i(m_{lf}-m_{li}-1)\phi} d\phi \end{aligned} \quad (2.1.9)$$

which means, to have a non-zero result, $\Delta m_l = \pm 1$. The y -polarised radiation gives the same result. \square

2.1.2 Fine structure

Spectroscopy measures transitions between energy levels, for example photoelectron spectroscopy measures that between electronic levels, and IR and Raman spectroscopy measures that between vibrational levels, and NMR between nuclear spin levels. By the same token, the energy differences between electronic spin levels can also be measured, and we'll develop some theoretical treatments to elucidate the fine structure of atomic spectra.

2.1.2.1 Orbital and spin magnetic moments

Orbital magnetic moment

The orbital magnetic moment can be derived from classical arguments: the current ($\frac{dq}{dt}$) of a circulating electron in the xy -plane at speed v at radius r is

$$I = -\frac{ev}{2\pi r} \quad (2.1.10)$$

The magnetic dipole in the z -direction is

$$m_z = IA = -\frac{ev}{2\pi r} \times \pi r^2 = -\frac{1}{2}evr \quad (2.1.11)$$

The z -component of the angular momentum is

$$l_z = m_e vr \quad (2.1.12)$$

so

$$m_z = -\underbrace{\frac{e}{2m_e}}_{\gamma_e} l_z \quad (2.1.13)$$

As the argument holds in all three cardinal directions, we arrive at the simple relation

$$\mathbf{m} = \gamma_e \mathbf{l} \quad (2.1.14)$$

where $\gamma_e = -e/2m_e$ is called the *gyromagnetic ratio* of the electron.

Now we're done with the classical and technically incorrect derivation, we need to introduce the quantum mechanical quantisation by claiming m_z is quantised like l_z , namely

$$m_z = \gamma_e m_l \hbar, \quad m_l = -l, -l+1, \dots, l \quad (2.1.15)$$

and we can define the *Bohr magneton* as $\mu_B = -\gamma_e \hbar = e\hbar/2m_e$ so we can write

$$m_z = -\mu_B m_l, \quad |m_l| = 0, 1, \dots, l \quad (2.1.16)$$

Spin magnetic moment

The orbital angular momentum has a classical analogue whereas the intrinsic spin does not. We can't rely on classical arguments here because there are none. It turns out that for the spin magnetic moment,

$$\begin{aligned} \mathbf{m} &= g_e \gamma_e \mathbf{s}, \quad g_e \approx 2 \\ m_z &= -g_e \mu_B m_s, \quad m_s = \pm \frac{1}{2} \end{aligned} \quad (2.1.17)$$

g_e is known as the *g-factor*². The nuclear spin magnetic moment obeys the same rule:

$$\begin{aligned} \mathbf{m} &= g_I \gamma_N \mathbf{I} \\ m_z &= g_I \mu_N m_I, \quad m_I = -I, -I+1, \dots, I \end{aligned} \quad (2.1.18)$$

where $\mu_N = e\hbar/2m_p$ is the *nuclear magneton*, and g_I is the nuclear magneton, of order 10^0 .

²This is twice the expected value from classical mechanics. The full derivation of the factor requires the relativistic Dirac equation and even then a much smaller (0.00232) correction requires quantum electrodynamics. One interesting way to think about this is that the quantum mechanical rotation operator will return $\text{Rot}(2\pi)|\chi\rangle = -|\chi\rangle$, which in essence says that a full rotation in quantum mechanics is 4π radian instead of 2π ! So the velocity of rotation and so on are all doubled.

Landé g -factor

For an electron with both orbital and spin magnetic moment, the resultant magnetic moment is given exactly³ as

$$\mathbf{m} = g_J \gamma \mathbf{J} \quad (2.1.19)$$

where g_J is the *Landé coefficient*, given by

$$g_J = \frac{3}{2} + \frac{S(S+1) - L(L+1)}{2J(J+1)} \quad (2.1.20)$$

2.1.2.2 Spin-orbit coupling

The Hamiltonian of spin-orbit coupling *in an isotropic electric field* can be derived exactly:

$$H_{\text{so}} = \xi(r) \mathbf{l} \cdot \mathbf{s} \quad (2.1.21)$$

where $\xi > 0$ and is proportional to the first derivative of the electric potential ϕ :

$$\xi(r) = -\frac{e}{2m_e^2 r c^2} \frac{d\phi}{dr} \quad (2.1.22)$$

and the *spin-orbit coupling constant* is given by the radial average of $\xi(r)$:

$$\hbar c \xi_{nl} = \langle n l m_l | \xi(r) | n l m_l \rangle \hbar^2 \quad (2.1.23)$$

Given the hydrogen-like Coulombic field $\phi = Ze/4\pi\epsilon_0 r$, the full Hamiltonian is then given as

$$H_{\text{so}} = \frac{Ze^2}{8\pi\epsilon_0 m_e^2 r^3 c^2} \mathbf{l} \cdot \mathbf{s} \quad (2.1.24)$$

we can derive

$$\xi_{nl} = \frac{\alpha^2 R Z^4}{n^3 l(l + \frac{1}{2})(l + 1)} \quad (2.1.25)$$

where $\alpha = e^2/4\pi\epsilon_0 \hbar c$ is the *fine-structure constant*. Note that it goes as Z^4 , which means the coupling constant rapidly increases in heavier elements.

The *Russell-Saunders coupling* scheme says about the energies

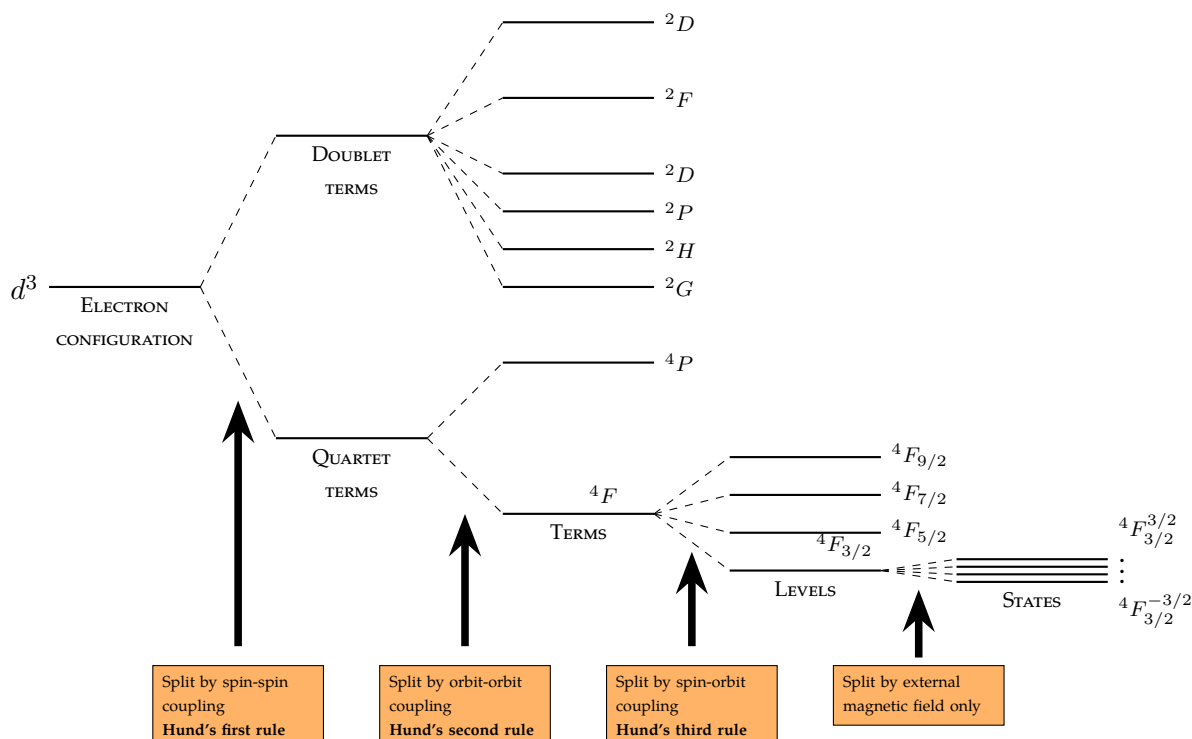
$$\text{spin-spin coupling} > \text{orbit-orbit coupling} > \text{spin-orbit coupling}$$

We need to explain this carefully:

³The only approximation being, $g_e = 2$ which goes into the 'offset' constant in front.

1. **Spin-spin coupling** means how strongly the spin angular momenta of the electrons interact, this is manifested by the Fermi holes, and more precisely, the *exchange interactions*. This splits a configuration into terms of different multiplicities. However this is rarely resolved on its own as the next level of branching, terms, are widely spaced and usually well resolved, as a consequence these energy levels don't get a name of their own.
2. **Orbit-orbit coupling** means how strongly orbital angular momenta interact - electrons that move in the same direction meet less than those that move in opposite directions. This splits the terms with the same multiplicity into terms.
3. Spin-orbit coupling was explained above. It splits terms into levels. However to fully explain Hund's third rule we need some further discussions, given below.
4. In the presence of external magnetic fields the levels are further split into states with the lowest M_J having the lowest energy.

So we can draw a schematic diagram like this



We now need to work out the exact expression for the energy separation between levels. Under the regime of Russell-Saunders coupling, we can apply first-order perturbation

theory (see Theorem 3.2.1.1) to get the energy correction:

$$E_{\text{so}} = \langle ls; jm_j | H_{\text{so}} | ls; jm_j \rangle = \langle ls; jm_j | \xi(r) \mathbf{l} \cdot \mathbf{s} | ls; jm_j \rangle \quad (2.1.26)$$

note that this is the *coupled picture*, which just specifies the individual angular momenta l and s and the resultant angular momentum j and its z -component m_j , of a single electron⁴, the same coupled picture, but applied to composite angular momenta, with l first building up to L and m to M , then both combine to give J , is the scheme under Russell-Saunders coupling. This discussion treats one-electron

We then write

$$\mathbf{j}^2 = |\mathbf{l} + \mathbf{s}|^2 = \mathbf{l}^2 + \mathbf{s}^2 + 2\mathbf{l} \cdot \mathbf{s} \quad (2.1.27)$$

so that

$$\begin{aligned} \mathbf{l} \cdot \mathbf{s} |ls; jm_j\rangle &= \frac{1}{2}(\mathbf{j}^2 - \mathbf{l}^2 - \mathbf{s}^2) |ls; jm_j\rangle \\ &= \frac{1}{2}\{j(j+1) - l(l+1) - s(s+1)\} |ls; jm_j\rangle \end{aligned} \quad (2.1.28)$$

where the boldface $\{\mathbf{l}, \mathbf{j}, \mathbf{s}\}$ are operators and the regular $\{l, j, s\}$ are the eigenvalues. We can write out the energy easily now, with reference to eq. 2.1.25,

$$\begin{aligned} E_{\text{so}} &= \frac{1}{2} \hbar^2 \{j(j+1) - l(l+1) - s(s+1)\} \langle ls; jm_j | \xi(r) | ls; jm_j \rangle \\ &= \frac{1}{2} \hbar c \xi_{nl} \{j(j+1) - l(l+1) - s(s+1)\} \\ &= Z^4 \alpha^2 \hbar c R \left\{ \frac{j(j+1) - l(l+1) - s(s+1)}{2n^3 l(l + \frac{1}{2})(l+1)} \right\} \end{aligned} \quad (2.1.29)$$

Ok, so far the discussion has been on single electron states, what about real atoms with many electrons? The spin-orbit coupling Hamiltonian is now approximately

$$H_{\text{so}} = \sum_{\text{outer shell}} \xi_i(r) \mathbf{l}_i \cdot \mathbf{s}_i \quad (2.1.30)$$

The outer shell label means that all inner shell electrons, *i.e.*, full and half shells, are expected to not have spin-orbit coupling because the sums of the dot product $\mathbf{l}_i \cdot \mathbf{s}_i$ is 0, since for a half-full shell, m_l runs from $+l$ to $-l$ while m_s is constant, and likewise for a full shell.

Therefore, for a shell less than half-full, say d^3 , Hund's first and second rules tell us that the orbitals will look like this:

⁴Note we've been using small case l, s, j , which is because so far we're only talking about single-electron states $|nlm_l\rangle$.

| | | | | | |
|-------|----|----|---|----|----|
| | ↑ | ↑ | ↑ | | |
| m_l | +2 | +1 | 0 | -1 | -2 |

The sum of the dot products will be positive, and referring to eq. 2.1.24, we note that, although the specific functional form will be different in actual many-electron atoms, in general $\xi(r) > 0$, so in our case of less than half-filled orbitals, the energy of interaction E_{so} will take on the form of eq. 2.1.29, and hence minimising J minimises the spin-orbit coupling energy.

Now consider a shell more than half-full, say d^6 , Hund's first and second rule still take precedence and tells us that the shell should be filled like this:

| | | | | | |
|-------|----|----|---|----|----|
| | ↑↓ | ↑ | ↑ | ↑ | ↑ |
| m_l | +2 | +1 | 0 | -1 | -2 |

Now the first five electrons will make no contribution to the spin-orbit coupling since their dot products add up to zero. The sixth electron has spin down, so the dot product is negative in sign, and this will make the energy of interaction negative of eq. 2.1.29, so now maximising J minimises energy.

You may think that, according to Hund's first and second rules which maximises M_S and M_L , which can never be negative, so \mathbf{S} and \mathbf{L} always point in the same direction, and so does their magnetic dipole moments (opposite), so what makes the interaction energy of the magnetic moments suddenly change its sign? Well, there's *no* physical law that says

$$H_{\text{so}} = \xi(r) \mathbf{L} \cdot \mathbf{S} \quad (2.1.31)$$

in other words,

$$H_{\text{so}} = \sum_{\text{outer shell}} \xi_i(r) \mathbf{l}_i \cdot \mathbf{s}_i \neq \xi(r) \mathbf{L} \cdot \mathbf{S} \quad (2.1.32)$$

because spin-orbit coupling is something that *each* electron experiences separately and differently. However, it is possible to write

$$H_{\text{so}} = \lambda \mathbf{L} \cdot \mathbf{S}, \quad \lambda = \pm \frac{1}{2S} \xi(r) \quad (2.1.33)$$

where the positive sign applies for a shell less than half full and negative for a shell more than half full, which is coherent with the above discussion. To see that, try d^7 , the sums of individual dot products gives $-3/2$, and eq. 2.1.33 also gives $-(3 \times \frac{3}{2}) / (2 \times \frac{3}{2}) = -3/2$.

2.2 MOLECULAR SPECTROSCOPY

2.2.1 Spectroscopic methods

The rotational spectra are associated with rotational fine structures that appear around vibrational lines in the IR spectra and they exclusively make up Raman bands. The branches are named as $X_{J''}$, where J'' is the lower J , and X is the ΔJ , which is $J_{\text{upper}} - J_{\text{lower}}$ ⁵.

| Letter | O | P | Q | R | S |
|------------|----|----|---|----|----|
| ΔJ | -2 | -1 | 0 | +1 | +2 |

We give a summary of the primary differences between the two spectroscopic methods below

| | Raman | IR |
|-------------------------------------|---|---|
| Sweeping incident frequency? | No, single beam of laser | Yes |
| What's being measured? | Frequencies of scattered light at right angles | Absorbance of incident lights |
| Structural requirement | Anisotropic polarisability | Anisotropic dipole moment |
| Rotational structure? | Only in rotational Raman spectra (low frequency exciting laser) | Always |
| Branches present? | QS in rotational, OQS in vibrational | PR; Q only in non-centrosymmetric molecules |

2.2.1.1 IR spectroscopy

Infrared spectroscopy uses sweeping infrared beam to excite the molecule to a higher vibrational state, and a sensor in the optical path of the incident beam measures the absorbance of the incident frequencies. As photons have $S = 1$, it must increase or decrease

⁵By absolute energy, so $v = 1, J = 2$ is higher than $v = 0, J = 6$

the J of the molecule by 1, therefore the rotational selection rule is $\Delta J = \pm 1$. This will be reflected as P and R branches in the rotational fine structures.

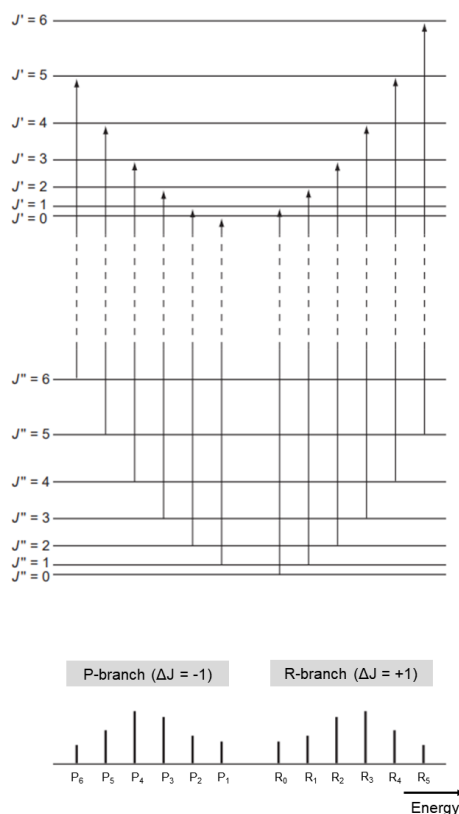


Figure 2.1: A schematic IR spectrum

The vibrational state must possess an anisotropic dipole moment wrt. the normal coordinate. Theorem 2.2.3.3 tells us whether a mode is IR-active.

2.2.1.2 Raman spectroscopy

Although complementary to IR spectroscopy, the theory of Raman spectroscopy is fundamentally different. An intense beam of laser at an appropriate, non-sweeping, monochromatic frequency is incident upon the sample, and the detector is placed at right angles to the incident beam to collect the scattered photons (note that no overall absorption of photons take place). The scattering can take place in three ways, we report $\Delta\nu$ and ΔJ as $\Delta\nu, \Delta J$:

1. **Rayleigh scattering:** Elastic scattering, no change in frequency of the photon, this

corresponds to unchanged rot-vibrational level and hence 0, 0.

2. **Stokes Raman scattering:** Inelastic scattering, reduced energy, corresponds to net excitation of rot-vibrational level, 0, +2 or +1, ± 2 .
3. **Anti-Stokes Raman scattering:** Inelastic scattering, increased energy, corresponds to net relaxation, 0, -2 or -1, ± 2 .

The $\Delta\nu = 0$ corresponds to *pure rotational Raman spectroscopy*, it uses low laser frequencies which do not have enough energy to excite the molecule to the next vibrational level, the $\Delta J = \pm 2$ correspond to the Stokes and anti-Stokes branches, with the anti-Stokes branch having the same intensities as many rotational levels are thermally accessible at room temperature, so it does not matter whether the initial state is lower or higher. A schematic pure Rotational Raman spectrum is shown in fig. 2.2 (the virtual levels are not shown)⁶:

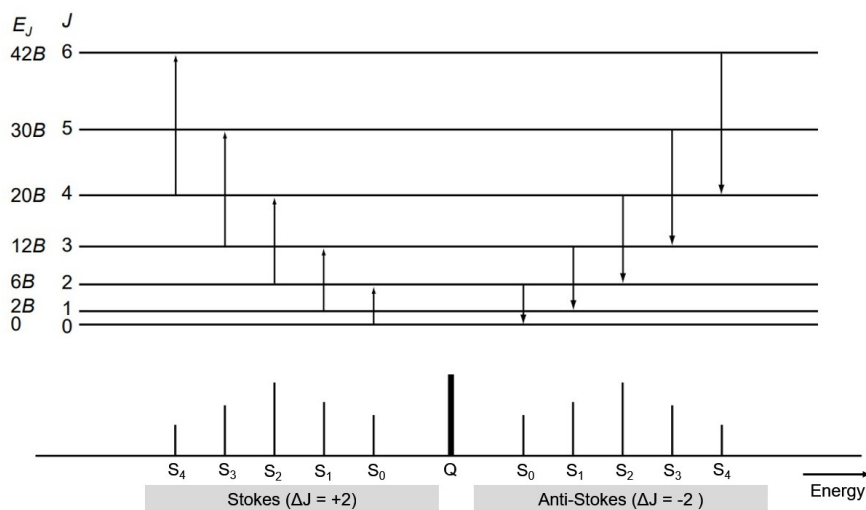


Figure 2.2: A schematic pure rotational Raman spectrum

The $\Delta\nu = \pm 1$ corresponds to *vibrational Raman spectroscopy*, which uses high laser frequencies that are enough to excite the molecules vibrationally. The $\Delta\nu = +1$ is the Stokes branch and $\Delta\nu = -1$ is the anti-Stokes. The $\Delta J = \pm 2$ should produce O- and S-branches as rotational fine structures but are too weak to observe in practice. A schematic vibrational Raman spectrum is shown in fig. 2.3:

⁶As the branch lettering is $J_{\text{upper}} - J_{\text{lower}}$, both branches are technically *S* in pure rotational Raman spectra.

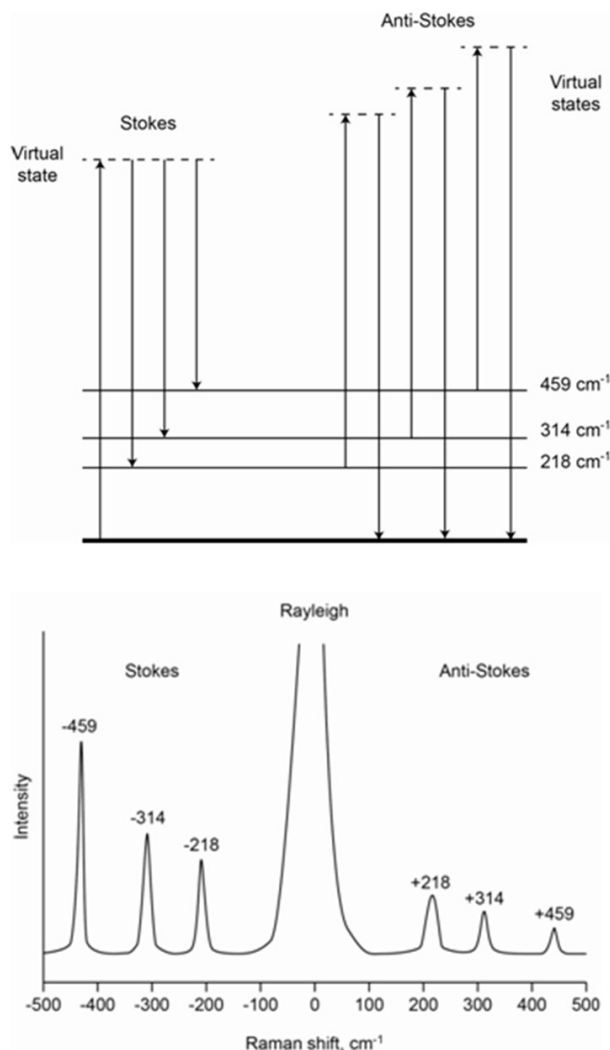


Figure 2.3: A schematic vibrational Raman spectrum

Whether a vibrational mode is active in vibrational Raman spectroscopy is given by Theorem 2.2.3.3 **not written yet, update with thrm for Raman.**

2.2.2 Transition frequencies

2.2.2.1 The vibrating rotor

The vibrational spectral line spacings are derived from the Morse potential, with energy levels

$$\epsilon_v = (v + \frac{1}{2})\tilde{\omega} - (v + \frac{1}{2})^2\tilde{\omega}x_e \quad (2.2.1)$$

The depth of the well, D_e , can be found by finding v_{\max} , *i.e.*, setting $d\epsilon_v/dv = 0$:

$$\begin{aligned} D_e &= \epsilon_{\max} - \epsilon_{\text{ZPE}} \\ &= \frac{\tilde{\omega}}{4x_e} = \frac{\hbar\omega}{4x_e} \end{aligned} \quad (2.2.2)$$

The bond dissociation energy can be found by taking the difference of the well depth and ground state energy:

$$\epsilon_B = \tilde{D}_e - \epsilon_0 = \tilde{D}_e(1 - x_e)^2 \quad (2.2.3)$$

The selection rule for the harmonic oscillator, $\Delta v = \pm 1$, proven in eq. 2.2.28, no longer hold true for the Morse oscillator, where we have $\Delta v = \pm 1, \pm 2, \dots$. However usually only the $0 \rightarrow 1$ (fundamental), $0 \rightarrow 2$ (first overtone) are strong and, with the temperature-dependent⁷ $1 \rightarrow 2$ (hot band) and the $0 \rightarrow 3$ (second overtone) very weak.

The transition frequencies are easy to calculate, some are listed here

| Transition | $0 \rightarrow 1$ | $0 \rightarrow 2$ | $0 \rightarrow 3$ | $1 \rightarrow 2$ |
|------------------|---------------------------------------|--|---|---------------------------------------|
| $\Delta\epsilon$ | $\tilde{\omega} - 2\tilde{\omega}x_e$ | $2\tilde{\omega} - 6\tilde{\omega}x_e$ | $3\tilde{\omega} - 12\tilde{\omega}x_e$ | $\tilde{\omega} - 4\tilde{\omega}x_e$ |

Because the vibrational wavefunction $\psi_n(x \equiv r - r_e)$ and the rotational wavefunction $Y_J^m(\theta, \phi)$ have different dependence, *i.e.*, degree of freedom. Therefore they are independent and we can write (we now adapt the notation in Footnote 1)

$$\begin{aligned} H\psi_{\text{rot-vib}} &\equiv (H_{\text{rot}} + H_{\text{vib}})Y_J^m\psi_v \\ &= (E_J + E_v)Y_J^m\psi_v \\ &= \left[BJ(J+1) + \hbar\omega \left(v + \frac{1}{2} \right) \right] \psi_{\text{rot-vib}} \end{aligned} \quad (2.2.4)$$

The convention is to write the energies in units of wavenumbers:

$$\tilde{E}_{v,J} = \tilde{B}J(J+1) + \tilde{v} \left(v + \frac{1}{2} \right), \quad (2.2.5)$$

where

$$\tilde{B} = \frac{\hbar}{4\pi\tilde{c}I}, \quad \tilde{v} = \frac{1}{2\pi\tilde{c}}\sqrt{\frac{k}{\mu}}. \quad (2.2.6)$$

$$\begin{aligned} \Delta v &= \pm 1 \text{ (overtones allowed but weak);} \\ \Delta J &= \pm 1. \end{aligned} \quad (2.2.7)$$

⁷The other bands are not temperature dependent because they all only involve transitions from the ground state which is always thermally accessible.

We have derived the vibrational spectral frequencies and the rotational frequencies as fine structures around the central vibrational frequencies. There's one problem we need to address, which is that the rotational constant \tilde{B} is not the same for each vibrational level because $\langle r \rangle$ increases as vibrational states increases⁸. We need to account for that in our derivation for exact frequencies in the rotational fine structure below:

$$\begin{aligned}\tilde{R}_{J''} &= \tilde{E}_{v=1, J'=J''+1} - \tilde{E}_{v=0, J''} \\ &= \left(v + \frac{3}{2}\right) \tilde{v} + \tilde{B}_1(J'' + 1)(J'' + 2) - \left(v + \frac{1}{2}\right) \tilde{v} - \tilde{B}_0 J''(J'' + 1) \\ &= \tilde{v}_0 + (\tilde{B}_1 - \tilde{B}_0)J''^2 + (3\tilde{B}_1 - \tilde{B}_0)J'' + 2\tilde{B}_1, \quad J'' = 0, 1, 2, \dots\end{aligned}\quad (2.2.8)$$

Likewise, the *P*-branch has

$$\tilde{P}_{J''} = \tilde{v} + (\tilde{B}_1 - \tilde{B}_0)J''^2 - (\tilde{B}_1 + \tilde{B}_0)J'', \quad J'' = 1, 2, 3, \dots, \quad (2.2.9)$$

where J'' cannot be zero as this implies J' starts at -1 .

In the approximation that $\tilde{B}_1 = \tilde{B}_0 \equiv \tilde{B}$, we arrive at

$$\begin{aligned}\tilde{R}_{J''} &= \tilde{v}_0 + 2\tilde{B}(J'' + 1) \\ \tilde{P}_{J''} &= \tilde{v}_0 - 2\tilde{B}J''.\end{aligned}\quad (2.2.10)$$

Property 2.2.2.1 (P and R branch peaks).

$$\begin{aligned}(\Delta J = +1) \tilde{R}_J &= \tilde{v}_0 + 2\tilde{B}_1 + (3\tilde{B}_1 - \tilde{B}_0)J + (\tilde{B}_1 - \tilde{B}_0)J^2, \quad J = 0, 1, 2, \dots \\ (\Delta J = -1) \tilde{P}_J &= \tilde{v}_0 - (\tilde{B}_1 + \tilde{B}_0)J + (\tilde{B}_1 - \tilde{B}_0)J^2, \quad J = 1, 2, 3, \dots\end{aligned}\quad (2.2.11)$$

Now, the equilibrium bond length increases as we go up in v with the Morse potential so I goes up and \tilde{B} goes down, so the term $(\tilde{B}_1 - \tilde{B}_0)J''^2$, although small, is always negative and increasingly so with increasing J'' . This means the *P* branch will be 'stretched' and the *R* branch will be 'compressed' more and more away from the central frequency.

2.2.2.2 Determination of rotational constants

The general expression for any J -transition from $v = 0$ to $v = 1$ is

$$\Delta\epsilon = \tilde{v}_0 + \tilde{B}_1 J'(J' + 1) - \tilde{B}_0 J''(J'' + 1). \quad (2.2.12)$$

⁸This is evident if we examine the functional form of the Morse potential - the oscillator spends more time on the right than the left as the vibrational states are ascended.

Considering peaks originating from the same J'' (lower J), this gives rise to two peaks, one in the P branch and one in the R branch, both labelled with the same J (conventionally labelled with lower J), and the difference is

$$\tilde{R}_J - \tilde{P}_J = 2\tilde{B}_1(2J + 1). \quad (2.2.13)$$

Similarly, for the same J' , we can find \tilde{B}_0 :

$$\tilde{R}_{J-1} - \tilde{P}_{J+1} = 2\tilde{B}_0(2J + 1) \quad (2.2.14)$$

Property 2.2.2.2 (Determination of rotational constants).

$$\begin{aligned} \tilde{R}_J - \tilde{P}_J &= 2\tilde{B}_1(2J + 1) \\ \tilde{R}_{J-1} - \tilde{P}_{J+1} &= 2\tilde{B}_0(2J + 1) \end{aligned} \quad (2.2.15)$$

Generally it is assumed that

$$\tilde{B}_v = \tilde{B}_e - \tilde{\alpha} \left(v + \frac{1}{2} \right), \quad (2.2.16)$$

where we need two \tilde{B}' s to determine the equilibrium rotational constant of a hypothetical vibrationless molecule. This can then be used to determine the true equilibrium bond length.

2.2.2.3 Q-branch

The selection rule $\Delta J = \pm 1$ for IR spectroscopy holds true strictly for *parallel vibrations* of linear molecules only, for molecules that have a perpendicular component of angular momentum, $\Delta J = 0$ is possible⁹. Examples include

- Any non-linear molecule, as there are other axes of rotation other than the principal axis, the angular momentum about which J is defined.
- Linear triatomic molecules, which have two degenerate bending modes, which create an oscillating dipole perpendicular to the principal axis, can combine to give rotational motions about the principal axis. This is the *vibrational angular momentum*.

⁹ J is defined as the angular momentum in the direction of the principal axis, note this is the intrinsic axis instead of the laboratory z -axis which defines m .

- Diatomics with electronic orbital angular momentum, like NO, whose ground state term symbol is $^2\Pi$, this angular momentum is intrinsic in origin, rather than induced by vibration

Since all quantum mechanical angular momenta are quantised in the same way (in steps of \hbar), a photon can excite these angular momenta. This means that conservation of angular momentum is now possible without a change in J ,

Property 2.2.2.3 (Selection rules for perpendicular vibrations).

$$\begin{aligned}\Delta J &= 0, \pm 1 \\ \Delta v &= \pm 1, \pm 2, \dots\end{aligned}\tag{2.2.17}$$

The Q -branch consists of all transitions with the same J , and the frequencies are given by

$$\begin{aligned}\epsilon' - \epsilon'' &= \left[\left(1 + \frac{1}{2}\right) \tilde{\omega} - \left(1 + \frac{1}{2}\right)^2 \tilde{\omega} x_e + \tilde{B}_1 J(J+1) \right] - \\ &\quad \left[\left(0 + \frac{1}{2}\right) \tilde{\omega} - \left(0 + \frac{1}{2}\right)^2 \tilde{\omega} x_e + \tilde{B}_0 J(J+1) \right] \\ &= \tilde{\omega} - 2\tilde{\omega} x_e + (\tilde{B}_1 - \tilde{B}_0) J(J+1) \\ &\equiv \tilde{\omega}_0 + (\tilde{B}_1 - \tilde{B}_0) J(J+1).\end{aligned}\tag{2.2.18}$$

Property 2.2.2.4 (Q -branch peaks). Q -branch peaks occur at

$$\tilde{\omega}_0 + (\tilde{B}_1 - \tilde{B}_0) J(J+1).\tag{2.2.19}$$

As $(\tilde{B}_1 - \tilde{B}_0)$ is small, and *positive* in this case as bending will result in smaller moment of inertia, this results in the spreading (towards higher wavenumbers) out of the Q -branch. This has effects on the P - and R -branches as well: now R branch will be spread out away from the central frequency and P will be compressed away from the central frequency, opposite of that in parallel stretches. All this is shown in fig. 2.4

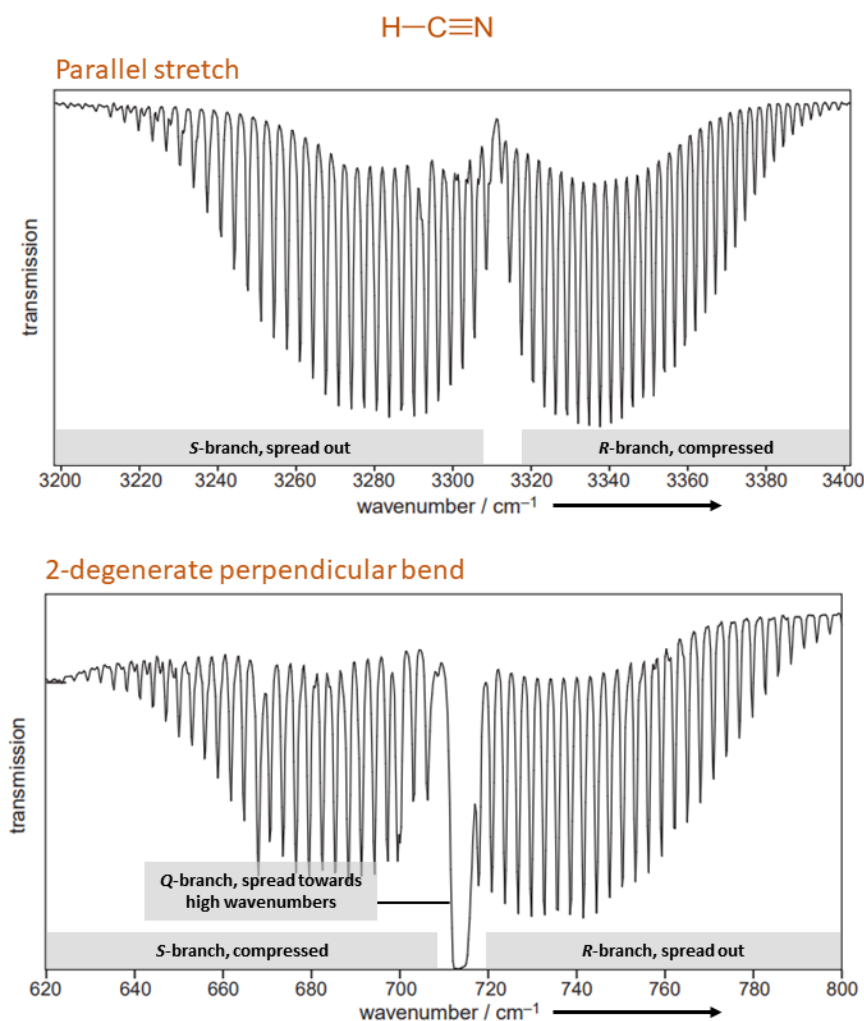


Figure 2.4: The IR spectrum of HCN, showing Q branch in only one of the vibrational modes.

The presence of Q -branch can be a diagnostic

Property 2.2.2.5 (Presence of Q -branch). The absence of Q -branch in any of the vibrational modes proves that the molecule has to be linear.

2.2.3 Polyatomic molecules

2.2.3.1 Normal coordinates

Theorem 2.2.3.1 (Normal coordinates). In any molecular vibration mode i it is always possible to define a normal coordinate Q_i that can replace the scaled coordinate q in the vibrational wavefunction.

Proof. The problem with polyatomic molecules is that they have $3N - 5$ (linear) or $3N - 6$ (non-linear) normal modes, therefore the same number of vibrational coordinates, q_i , and we can write the change in potential energy as

$$\begin{aligned} dV &= \frac{1}{2} \sum_{i=1}^{N_{\text{vib}}} \sum_{j=1}^{N_{\text{vib}}} \frac{\partial^2 V}{\partial q_i \partial q_j} q_i q_j + (\text{higher derivatives}) \\ &= \frac{1}{2} \sum_{i=1}^{N_{\text{vib}}} \sum_{j=1}^{N_{\text{vib}}} f_{ij} q_i q_j + \dots \end{aligned} \quad (2.2.20)$$

Unfortunately, the cross-terms makes the Schrödinger equation intractable. A theorem from classical mechanics can transform the cross terms into *normal coordinates*:

$$dV \approx \frac{1}{2} \sum_{j=1}^{N_{\text{vib}}} F_j Q_j^2 \quad (2.2.21)$$

□

Property 2.2.3.1 (Symmetry of normal coordinates). The normal coordinate Q_i has the same symmetry as the normal mode.

Proof. The normal coordinate Q_i defines the normal mode as the linear combination of vibrational coordinates that make up the normal mode. Therefore *must* transform as the symmetry species of the normal mode. This pointing out the chicken-and-egg situation rather than a proof. □

2.2.3.2 Symmetry of vibrational wavefunctions

As the transition probability involves evaluating integrals of the form $\langle \psi_i | \mu | \psi_j \rangle$, we need to work out the symmetry species spanned by the vibrational wavefunctions.

Theorem 2.2.3.2 (Symmetry of vibrational wavefunctions). For harmonic oscillators, odd state vibrational wavefunctions always transform as the totally symmetric IR.

Even state vibrational wavefunctions always transform as the IR of the normal mode.

For degenerate normal modes, the first excited state transforms as the normal mode, but higher excited states are more complex **what?**

Proof. As the unnormalised ground state is

$$\psi_0 = e^{-\frac{1}{2}Q_i^2} \quad (2.2.22)$$

For a non-degenerate irrep, this must return characters as 1's, *i.e.*, the totally symmetric IR.

The first excited state is

$$\psi_1 = Q_i e^{-\frac{1}{2}Q_i^2}, \quad (2.2.23)$$

we know the exponential bit transforms as the tot. sym. IR, and that Q_i transforms as the normal mode (Theorem 2.2.3.1), therefore ψ_1 transforms as the symmetry species of the normal mode.

The form of the Hermite polynomials **is this true? specific only to Hermite?** dictates that even-numbered functions contains only even exponents and constants, and that odd-numbered functions contains only odd exponents and no constants, *i.e.*, factorisable to $[Q_i(\text{tot. sym. function})]$, which means they all transform as the normal mode. \square

2.2.3.3 Symmetry selection rules for infrared

For transition between energy levels of a single normal mode, we assume we start at the ground state, we need to know whether the integral

$$R_{v_i, v'_i} = \langle \psi_{v'_i} | \mu | \psi_{v_i} \rangle \quad (2.2.24)$$

is non-zero, where μ transforms like the principle Cartesian axes:

Definition 2.2.3.1 (Dipole moment operator). As the dipole moment μ is only dependent on the normal coordinate Q_i , we can write

$$\mu = \mu_0 + \left(\frac{d\mu}{dQ_i} \right)_{Q_i=0} Q_i + \frac{1}{2} \left(\frac{d^2\mu}{d(Q_i)^2} \right)_{Q_i=0} Q_i^2 + \dots \quad (2.2.25)$$

If we make the approximation that the partial charges are constant *i.e.*, dipole moment varies linearly with normal coordinate, with the first term vanished by orthonormality, we are left with the definition of the dipole moment:

$$\mu \equiv \left(\frac{d\mu}{dx} \right)_0 Q_i \quad (2.2.26)$$

Invoking theorem **which one**, we arrive at the result

Theorem 2.2.3.3 (Transition rules). For a fundamental transition (transitions from the ground state energy of a normal mode to an excited state of the same mode) to be allowed in infrared, the IR of the normal mode has to be the same as that of x, y or z.

However symmetry allowed transitions can still have their integrals evaluated to zero by functional forms determined not by symmetry. For example, the selection rule for the harmonic oscillator is $\nu = \pm 1$, this is derived from the ladder operators:

As stated above, we can always write the normal coordinate as

$$Q_i = \left(\frac{\hbar}{2\mu\omega} \right)^{1/2} (a_- + a_+) \quad (2.2.27)$$

where the ladder operators also have their scaled coordinate replaced by the normal coordinate. It is now easy to write

$$\begin{aligned} \langle v_i | \mu | v_j \rangle &\propto \langle v_i | (a_- + a_+) | v_j \rangle \\ &= \delta_{i,j-1} + \delta_{i,j+1} \end{aligned} \quad (2.2.28)$$

and it is apparent that only when $j = i \pm 1$, the matrix element is nonzero.

2.2.4 Electronic spectroscopy

2.2.4.1 The Franck-Condon principle

The Einstein B coefficient for induced absorption or emission¹⁰ is given as

$$B_{mn} = \frac{1}{3\epsilon_0\hbar^2} |\langle \psi_m | \hat{\mu} | \psi_n \rangle|^2 \quad (2.2.29)$$

Under the Born-Oppenheimer approximation, ignoring the rotational changes, as they are too fine to be resolved, we can write

$$\psi_i = \psi_{i,E} \psi_{i,V}, \quad \hat{\mu} = \hat{\mu}_E + \hat{\mu}_V \quad (2.2.30)$$

¹⁰The A coefficient is for spontaneous emission.

The transition dipole moment is now

$$\begin{aligned}
 & |\langle \psi'' | \hat{\mu} | \psi' \rangle|^2 \\
 &= |\langle \psi''_E \psi''_V | \hat{\mu}_E + \hat{\mu}_V | \psi'_E \psi'_V \rangle|^2 \\
 &= |\langle \psi''_E \psi''_V | \hat{\mu}_E | \psi'_E \psi'_V \rangle + \langle \psi''_E \psi''_V | \hat{\mu}_V | \psi'_E \psi'_V \rangle|^2 \\
 &= |\langle \psi''_E | \hat{\mu}_E | \psi'_E \rangle \langle \psi''_V | \psi'_V \rangle + \langle \psi''_E | \psi'_E \rangle \langle \psi''_V | \hat{\mu}_V | \psi'_V \rangle|^2 \\
 &= |\langle \psi''_E | \hat{\mu}_E | \psi'_E \rangle|^2 |\langle \psi''_V | \psi'_V \rangle|^2
 \end{aligned} \tag{2.2.31}$$

The electronic wavefunctions are orthonormal, but not the vibrational states *belonging to different electronic states* as they are on different Morse potentials and are solutions to different Schrödinger equations.

The first bit is the electronic transition dipole, and group theory can tell us whether that vanishes. The second part is known as the *Franck-Condon factor*, which tells us the extent of overlap between two vibrational wavefunctions. The *Franck-Condon principle* is the assumption that there's no change in bond length before and after the electronic transition *i.e.*, the overlap F-C factor can tell us the probability of transition.

Theorem 2.2.4.1 (The Franck-Condon principle). Because the nuclei are much more massive than electrons, an *electronic* transition takes place while the nuclei in a molecule are effectively stationary. The transition probability between two vibrational states in two electronic states is given by $|\langle \psi''_V | \psi'_V \rangle|^2$.

This will give progressions that start from the lower vibrational states $v'' = 0, 1, 2, \dots$, which each successive progression at lower wavenumbers and lower intensities as the states become less populated.

The transition energies can be represented as follows

$$\begin{aligned}
 \tilde{\nu}' - \tilde{\nu}'' &= (\epsilon'_0 - \epsilon''_0) + \{[(v' + \frac{1}{2})\tilde{\omega}'_e - \tilde{\omega}'_e x'_e(v' + \frac{1}{2})^2] - [(v'' + \frac{1}{2})\tilde{\omega}''_e - \tilde{\omega}''_e x''_e(v'' + \frac{1}{2})^2]\} \\
 &\equiv T_e + \Delta\tilde{E}_v
 \end{aligned} \tag{2.2.32}$$

This allows us to compute parameters such as the equilibrium bond length and the dissociation energy of an excited electronic mode.

Chapter 3

Approximation methods

The Hamiltonian for a general neutral atom with nuclear charge Ze is

$$H = \underbrace{\sum_{j=1}^Z \left[-\frac{\hbar^2}{2m} \nabla_j^2 - \left(\frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r_j} \right) \right]}_{\text{kinetic energy+nuclear attraction}} + \underbrace{\frac{1}{2} \left(\frac{1}{4\pi\epsilon_0} \right) \sum_j^Z \sum_{k \neq j}^Z \frac{e^2}{|\mathbf{r}_j - \mathbf{r}_k|}}_{\text{interelectronic repulsion}}. \quad (3.0.1)$$

The factor of $1/2$ merely corrects for the fact that the double summation counts over each pair twice. We then wish to solve for

$$H\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_Z) = E\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_Z). \quad (3.0.2)$$

And the only acceptable solutions according to the symmetrisation requirement is when, coupled with the spin, the complete wavefunction is antisymmetric wrt exchange of electrons.

This equation cannot be solved exactly and we must resort to approximation methods which will be discussed in this chapter.

3.1 THE VARIATIONAL METHOD

3.1.1 Theory

Consider now the ground state of an arbitrary system. The ground state Schrödinger equation states

$$H|\psi_0\rangle = E_0|\psi_0\rangle. \quad (3.1.1)$$

Right-multiplying by $\langle\psi_0|$ we have

$$E_0 = \frac{\langle\psi_0|H|\psi_0\rangle}{\langle\psi_0|\psi_0\rangle} \quad (3.1.2)$$

We introduce the theorem

Theorem 3.1.1.1 (The variational principle). For *any other* arbitrary function ϕ with the same arguments (provided that the integral exists of course), the corresponding energy

$$E_\phi = \frac{\langle\phi|H|\phi\rangle}{\langle\phi|\phi\rangle} \quad (3.1.3)$$

must have that

$$E_\phi \geq E_0, \quad (3.1.4)$$

where the equality holds only for $\phi = \psi_0$.

Proof. Because the set of eigenfunctions of H is complete, we can express *any* function ϕ as

$$|\phi\rangle = \sum_{n=0}^{\infty} c_n |\psi_n\rangle. \quad (3.1.5)$$

Where, from the orthonormality of ψ_n 's,

$$c_n = \langle\psi_n|\phi\rangle. \quad (3.1.6)$$

So Equation 3.1.3 now reads

$$E_\phi = \frac{\langle\phi|H|\phi\rangle}{\langle\phi|\phi\rangle} = \frac{c_n^* c_n \langle\sum \psi_n | H | \sum \psi_n\rangle}{c_n^* c_n \langle\sum \psi_n | \sum \psi_n\rangle} = \frac{\sum |c_n|^2 E_n}{\sum |c_n|^2}. \quad (3.1.7)$$

Subtracting E_0 from both sides, we have

$$E_\phi - E_0 = \frac{\sum |c_n|^2 (E_n - E_0)}{\sum |c_n|^2}, \quad (3.1.8)$$

which *must* be non-negative as E_n is by definition greater than E_0 , and all $|c_n|^2$ are non-negative as well. And we observe that the equality holds necessarily when $\phi = \psi_0$ since all non-zero c_n 's must be 0, but at $n = 0$, $E_n = E_0$ also, so the entire summation in the numerator vanishes for all n . \square

The variational principle says that we can get an upper bound to E_0 by using any trial function as we wish, with closer approximation if ϕ is functionally more similar to ψ_0 . In general, we will choose a trial function that depends on some arbitrary *variational parameters*, $\alpha, \beta, \gamma, \dots$. The energy then will also depend upon these parameters, *i.e.*,

$$E_\phi(\alpha, \beta, \gamma, \dots) \geq E_0. \quad (3.1.9)$$

Now we can minimise E_ϕ wrt each of the parameters, so the best possible ground-state energy of a specific trial wavefunction can be found.

3.1.2 Examples

Hydrogen ground state

As an example, we try to approximate the hydrogen ground state wavefunction. We just need to know the Hamiltonian:

$$H = \frac{\hbar^2}{2m} \frac{d}{dr} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{e^2}{4\pi\epsilon_0 r}. \quad (3.1.10)$$

Say we guess that the wavefunction is a Gaussian (and we know from empirical evidence that it will not depend on θ and ϕ):

$$\phi(r) = e^{-\alpha r^2}. \quad (3.1.11)$$

Let's now calculate E_ϕ .

$$\langle \phi | H | \phi \rangle = 4\pi \int_0^\infty \phi^*(r) H \phi(r) r^2 dr = \frac{3\hbar^2 \pi^{3/2}}{4\sqrt{2m}\alpha^{1/2}} - \frac{e^2}{4\epsilon\alpha}, \quad (3.1.12)$$

where we have used the standard integrals in Property 1.2.3.5. Similarly,

$$\langle \phi | \phi \rangle = 4\pi \int_0^\infty \phi^* \phi r^2 dr = \left(\frac{\pi}{2\alpha} \right)^{3/2} \quad (3.1.13)$$

where we note that in general, these trial functions are, of course, not normalised, but nowhere in the derivation of the variational principal required that the wavefunctions,

trial or actual, be normalised.

We can therefore compute the trial energy

$$E(\alpha) = \frac{3\hbar^2\alpha}{2m} - \frac{e^2\alpha^{1/2}}{\sqrt{2\epsilon_0\pi^{3/2}}}. \quad (3.1.14)$$

Setting the first derivative to zero we obtain

$$\alpha_{\min} = \frac{m^2 e^4}{18\pi^3 \epsilon_0^2 \hbar^4}, \quad (3.1.15)$$

so

$$E_{\min} = -\frac{4}{3\pi} \left(\frac{me^4}{16\pi^2 \epsilon_0^2 \hbar^2} \right) \approx -0.424 \left(\frac{me^4}{16\pi^2 \epsilon_0^2 \hbar^2} \right), \quad (3.1.16)$$

where the exact value is

$$E_0 = -0.500 \left(\frac{me^4}{16\pi^2 \epsilon_0^2 \hbar^2} \right). \quad (3.1.17)$$

Pretty good approximation despite the wrong choice of trial function. And sure enough, by using the trial function of

$$\phi = e^{-\alpha r} \quad (3.1.18)$$

we can get the exact ground state energy and wavefunction.

Helium ground state

The previous example was based on a system we do know how to solve exactly, and now we apply it to one that we don't: the helium atom, with the Hamiltonian

$$\begin{aligned} H &= -\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2) - \frac{e^2}{2\pi\epsilon} \left(\frac{1}{r_1} + \frac{1}{r_2} \right) + \frac{e^2}{4\pi\epsilon_0} \frac{1}{r_{12}} \\ &= H_1 + H_2 + \frac{e^2}{4\pi\epsilon_0} \frac{1}{r_{12}}, \end{aligned} \quad (3.1.19)$$

where H_i 's are the Hamiltonian for a single electron atom with $Z = 2$, and they satisfy the equation

$$H_i \psi(r_i, \theta_i, \phi_i) = E_i \psi(r_i, \theta_i, \phi_i). \quad (3.1.20)$$

We choose our trial function by ignoring the interelectronic repulsion term, *i.e.*,

$$\phi_0(\mathbf{r}_1, \mathbf{r}_2) = \psi_{1s}(\mathbf{r}_1)\psi_{1s}(\mathbf{r}_2), \quad (3.1.21)$$

where ψ_{1s} is the $Z = 2$ ground state hydrogenlike wavefunction:

$$\psi_{1s}(\mathbf{r}_i) = \left(\frac{Z^3}{\pi a^3} \right)^{1/2} e^{-Zr_i/a}, \quad i = 1, 2. \quad (3.1.22)$$

So our trial function is

$$\phi(\mathbf{r}_1, \mathbf{r}_2) = \frac{Z^3}{a^3\pi} e^{-Z(r_1+r_2)/a}. \quad (3.1.23)$$

As the trial function is normalised this time, we only have to evaluate the numerator:

$$E(Z) = \iint \phi_0^*(\mathbf{r}_1, \mathbf{r}_2) H \phi_0(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2. \quad (3.1.24)$$

This integral can seem daunting but we can greatly simplify it by realising ψ is an eigenfunction of the hydrogenlike Hamiltonian:

$$H\psi = -\frac{Z^2\hbar^2}{2ma^2}\psi. \quad (3.1.25)$$

We now re-write the Helium Hamiltonian in terms of hydrogenlike Hamiltonians, keeping in mind that Z is not necessarily 2 now as it is a variational parameter:

$$\begin{aligned} H &= -\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2) - \frac{e^2}{2\pi\epsilon} \left(\frac{1}{r_1} + \frac{1}{r_2} \right) + \frac{e^2}{4\pi\epsilon_0} \frac{1}{r_{12}} \\ &= -\frac{\hbar^2}{2m}\nabla_1^2 - \frac{Ze^2}{4\pi\epsilon_0 r_1} - \frac{\hbar^2}{2m}\nabla_2^2 - \frac{Ze^2}{4\pi\epsilon_0 r_2} + \frac{(Z-2)e^2}{4\pi\epsilon_0 r_1} + \frac{(Z-2)e^2}{4\pi\epsilon_0 r_2} + \frac{e^2}{4\pi\epsilon_0 r_{12}} \\ &= H_{Z1} + H_{Z2} + \frac{(Z-2)e^2}{4\pi\epsilon_0 r_1} + \frac{(Z-2)e^2}{4\pi\epsilon_0 r_2} + \frac{e^2}{4\pi\epsilon_0 r_{12}}. \end{aligned} \quad (3.1.26)$$

So

$$\begin{aligned} E(Z) &= \iint \phi \left[-\frac{Z^2\hbar^2}{2ma^2} - \frac{Z^2\hbar^2}{2ma^2} + \frac{(Z-2)e^2}{4\pi\epsilon_0 r_1} + \frac{(Z-2)e^2}{4\pi\epsilon_0 r_2} + \frac{e^2}{4\pi\epsilon_0 r_{12}} \right] \phi d\mathbf{r}_1 d\mathbf{r}_2 \\ &= \frac{Z^6}{a^6\pi^2} 16\pi^2 \iint \left[-\frac{Z^2\hbar^2}{ma^2} + \frac{(Z-2)e^2}{4\pi\epsilon_0 r_1} + \frac{(Z-2)e^2}{4\pi\epsilon_0 r_2} \right] e^{-2Z(r_1+r_2)/a} r_1^2 r_2^2 d\mathbf{r}_1 d\mathbf{r}_2 \\ &\quad + \frac{Z^6}{a^6\pi^2} \iint \frac{e^2}{4\pi\epsilon_0 r_{12}} e^{-2Z(r_1+r_2)/a} d\mathbf{r}_1 d\mathbf{r}_2 \\ &= \frac{16Z^6}{a^6} \iint \left[-\frac{Z^2\hbar^2}{4\pi\epsilon_0 a} + \frac{(Z-2)e^2}{4\pi\epsilon_0 r_1} + \frac{(Z-2)e^2}{4\pi\epsilon_0 r_2} \right] e^{-2Z(r_1+r_2)/a} r_1^2 r_2^2 d\mathbf{r}_1 d\mathbf{r}_2 \\ &\quad + \frac{Z^6}{a^6\pi^2} \iint \frac{e^2}{4\pi\epsilon_0 r_{12}} e^{-2Z(r_1+r_2)/a} d\mathbf{r}_1 d\mathbf{r}_2 \end{aligned} \quad (3.1.27)$$

We have isolated the integral involving r_{12} because they are implicitly dependent on θ . There are four integrals to solve, for the first three we use the standard integral Equa-

tion 1.2.106a:

$$\begin{aligned}
 & \frac{16Z^6}{a^6} \iint -\frac{Z^2}{4\pi\epsilon_0 a} e^{-2Z(r_1+r_2)} r_1^2 r_2^2 dr_1 dr_2 \\
 &= -\frac{4Z^8 e^2}{\pi\epsilon_0 a^7} \left(\int e^{-2Zr/a} r^2 dr \right)^2 \\
 &= -\frac{4Z^8 e^2}{\pi\epsilon_0 a^7} \left(\frac{a^3}{4Z^3} \right)^2 \\
 &= -\frac{Z^2 e^2}{4\pi\epsilon_0 a} \\
 &= -Z^2 \left(\frac{me^4}{16\pi^2\epsilon_0^2 \hbar^2} \right) = -Z^2 \text{ Hartree,}
 \end{aligned} \tag{3.1.28}$$

and the second and third are the same integral:

$$\begin{aligned}
 & \frac{16Z^6(Z-2)e^2}{4\pi\epsilon_0 a^6} \iint r_1 r_2^2 e^{-2Z(r_1+r_2)/a} dr_1 dr_2 \\
 &= \frac{16Z^6(Z-2)e^2}{4\pi\epsilon_0 a^6} \frac{a^5}{16Z^5} \\
 &= Z(Z-2) \text{ Ha.}
 \end{aligned} \tag{3.1.29}$$

The last integral is trickier, we first need to choose the z -axis to be one of the vectors, say r_1 , so we can use our usual spherical coordinates:

$$r_{12} = (r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta)^{1/2}. \tag{3.1.30}$$

The integral becomes

$$\begin{aligned}
 & \frac{Z^6 e^2}{4\pi^3 \epsilon_0 a^6} \iint \frac{1}{(r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta)^{1/2}} e^{-2Z(r_1+r_2)/a} d\mathbf{r}_1 d\mathbf{r}_2 \\
 &= \frac{Z^6 e^2}{4\pi^3 \epsilon_0 a^6} 4\pi^2 \int_0^\pi \sin \theta_2 d\theta_2 \int_0^\infty \int_0^\infty \int_0^\pi e^{-2Z(r_1+r_2)/a} r_1^2 r_2^2 \frac{\sin \theta_1 d\theta_1 dr_1 dr_2}{(r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta_1)^{1/2}} \\
 &= \frac{2Z^6 e^2}{\pi\epsilon_0 a^6} \int_0^\infty \int_0^\infty e^{-2Z(r_1+r_2)/a} r_1 r_2 (r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta_1)^{1/2} \Big|_0^\pi dr_1 dr_2 \\
 &= \frac{2Z^6 e^2}{\pi\epsilon_0 a^6} \int_0^\infty \int_0^\infty e^{-2Z(r_1+r_2)/a} r_1 r_2 \left[(r_1^2 + r_2^2 + 2r_1 r_2)^{1/2} - (r_1^2 + r_2^2 - 2r_1 r_2)^{1/2} \right] dr_1 dr_2.
 \end{aligned} \tag{3.1.31}$$

Now, the terms in the square brackets must be positive as we are computing energy. When $r_2 < r_1$ we have the terms equal to $2r_2$, and when $r_2 > r_1$ they are equal to $2r_1$. Making

r_1 run from 0 to infinity and splitting the integral over r_2 into two parts, the integral is

$$\begin{aligned}
 & \frac{2Z^6 e^2}{\pi \epsilon_0 a^6} \int_0^\infty r_1 e^{-2Zr_1/a} dr_1 \left(2 \int_0^{r_1} r_2^2 e^{-2Zr_2/a} dr_2 + 2r_1 \int_{r_1}^\infty r_2 e^{-2Zr_2/a} dr_2 \right) \\
 &= \frac{4Z^6 e^2}{\pi \epsilon_0 a^6} \int_0^\infty r_1 e^{-\alpha r_1} \left[\frac{2}{\alpha^3} - \left(\frac{r_1}{\alpha^2} + \frac{2}{\alpha^3} \right) e^{-\alpha r_1} \right] dr_1 \\
 &= \frac{4Z^6 e^2}{\pi \epsilon_0 a^6} \left(\frac{2}{\alpha^5} - \frac{1}{4\alpha^5} - \frac{1}{2\alpha^5} \right) \\
 &= \frac{5Z}{8} \text{ Ha.}
 \end{aligned} \tag{3.1.32}$$

Finally we can assemble the entire integral:

$$E(Z) = Z^2 - \frac{27}{8}Z \tag{3.1.33}$$

in Hartrees, minimising we have

$$Z_{\min} = \frac{27}{16} \Rightarrow E_{\min} \approx -2.8477. \tag{3.1.34}$$

The experimental value is -2.9033 , we have reached a pretty good approximation even when we have disregarded interelectronic repulsion completely. The variational principal in a way compensates that by returning us a Z less than 2, which can be understood as an **effective nuclear charge**.

3.1.3 Secular determinants

A trial function with more than one parameter can return very good results, as to be expected since the ground state wavefunction can be expressed as an infinite series of any functions, the more terms we include, the better we can approximate. We use a trial function of the form

$$\phi = \sum_{n=1}^N c_n f_n, \tag{3.1.35}$$

where c_n are variational parameters and f_n are arbitrarily chosen functions. We restrict ourselves to two terms, and c_n and f_n can be complex

$$\phi = c_1 f_1 + c_2 f_2. \tag{3.1.36}$$

To get E_ϕ , we compute

$$\begin{aligned}\langle \phi | H \phi \rangle &= \langle (c_1 f_1 + c_2 f_2) | H (c_1 f_1 + c_2 f_2) \rangle \\ &= |c_1|^2 H_{11} + c_1^* c_2 H_{12} + c_1 c_2^* H_{21} + |c_2|^2 H_{22},\end{aligned}\quad (3.1.37)$$

where

$$H_{ij} = \langle f_i | H f_j \rangle. \quad (3.1.38)$$

Similarly,

$$\langle \phi | \phi \rangle = |c_1|^2 S_{11} + c_1^* c_2 S_{12} + c_1 c_2^* S_{21} + |c_2|^2 S_{22}. \quad (3.1.39)$$

where

$$S_{ij} = \langle f_i | f_j \rangle. \quad (3.1.40)$$

So we can write (to facilitate implicit differentiation)

$$|c_1|^2 H_{11} + c_1^* c_2 H_{12} + c_1 c_2^* H_{21} + |c_2|^2 H_{22} = E_\phi (|c_1|^2 S_{11} + c_1^* c_2 S_{12} + c_1 c_2^* S_{21} + |c_2|^2 S_{22}). \quad (3.1.41)$$

We now minimise $E(c_1, c_2)$. However, we want the final expression to be in terms of c_1 and c_2 , not their conjugates, purely for the sake of brevity, so we instead minimise c_1^* and c_2^* :

$$\frac{\partial E}{\partial c_1^*} = 0 \Rightarrow c_1(H_{11} - E_\phi S_{11}) + c_2(H_{12} - E_\phi S_{12}) = 0, \quad (3.1.42)$$

and minimising c_2^* we have

$$c_1(H_{21} - E_\phi S_{21}) + c_2(H_{22} - E_\phi S_{22}) = 0. \quad (3.1.43)$$

where we need the property, *ad hoc*, that

Property 3.1.3.1 (Complex differentiation).

$$\frac{\partial f}{\partial z} = \frac{1}{2} \left(\frac{\partial f}{\partial x} - i \frac{\partial f}{\partial y} \right), \quad \frac{\partial f}{\partial z^*} = \frac{1}{2} \left(\frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} \right). \quad (3.1.44)$$

Proof. This is a heuristic proof and is not rigorous. See Cauchy-Riemann equation for theoretical foundation. Let

$$z \equiv x + iy, \Rightarrow z^* = x - iy, \quad (3.1.45)$$

so

$$x = \frac{1}{2}(z + z^*), \quad y = -\frac{1}{2}i(z - z^*). \quad (3.1.46)$$

So

$$\frac{\partial f}{\partial z} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial z} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial z} = \frac{1}{2} \left(\frac{\partial f}{\partial x} - i \frac{\partial f}{\partial y} \right), \quad (3.1.47)$$

and so on. □

Anyway, casting the above equations in matrix form, we have

$$\begin{pmatrix} H_{11} - E_{\phi} S_{11} & H_{12} - E_{\phi} S_{12} \\ H_{21} - E_{\phi} S_{21} & H_{22} - E_{\phi} S_{22} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = 0, \quad (3.1.48)$$

which means there will only be non-trivial solutions for c_1 and c_2 when the determinant

$$\begin{vmatrix} H_{11} - E_{\phi} S_{11} & H_{12} - E_{\phi} S_{12} \\ H_{21} - E_{\phi} S_{21} & H_{22} - E_{\phi} S_{22} \end{vmatrix} = 0. \quad (3.1.49)$$

Determinants like this are called **secular determinants**. Expanding and solving them will give us a quadratic equation in E_{ϕ} without explicitly computing c_n . We take the smaller of the two solutions as the upper bound estimate. Once we have determined E_{ϕ} we can invert the matrix equation to find the c_n 's and subsequently normalise the approximate solution. This can be trivially generalised to trial functions of N terms.

Just as an aside, the f_n 's themselves can include variational parameters to obtain even better approximations, but however we can't use secular determinants in that case and numerical computations are required.

3.1.4 The Hückel approximations

To compute the secular determinant more quickly, the Hückel approximations can be used. The approximations are

1. The overlap integral $S_{ij} = 0$ for $i \neq j$
2. The Hamiltonian matrix element is nonzero only for neighbouring atoms.

A number of simplified notations are adopted:

1. $H_{ii} \equiv \alpha_{ii} \equiv \alpha$
2. $H_{ij} \equiv \beta_{ij} = \beta_{ji}$

The method proceeds as follows

1. Find out the point group the molecule belongs to
2. Identify unique sets of orbitals that are interconverted, and discard those belonging to the same symmetry species, *i.e.*, they are never interconverted
3. Reduce the representation
4. Set up symmetry orbitals for all sets of orbitals
5. Solve the secular equation for symmetry orbitals belonging to the same symmetry to obtain energies and coefficients

This is a long list, and it best illustrated by examples

3.1.4.1 Example 1

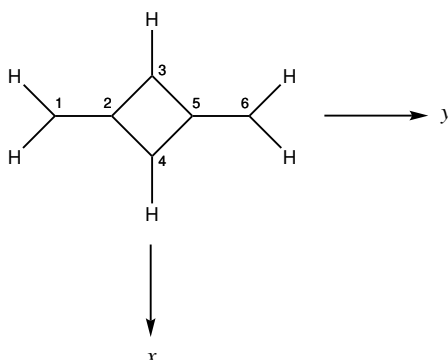


Figure 3.1: All the carbon atoms are sp^2 hybridised with the p orbitals perpendicular to the plane of paper. Only σ framework is shown on the plane of paper.

Point group and basis

The molecule belongs to D_{2h} point group. It is clear that the p -orbital pairs on atoms 1 and 6, and 2 and 5 are equivalent under operations of D_{2h} , so we can discard one of them since that makes reducing representations easier. The pairs 2 and 5 and 3 and 4 are not mixed or interconverted, so with them each setting up a basis, we can write

$$\begin{aligned}\Gamma_1 &= \phi_2 + \phi_5 = B_{3g} \oplus B_{1u} \\ \Gamma_2 &= \phi_3 + \phi_4 = B_{2g} \oplus B_{1u}\end{aligned}\tag{3.1.50}$$

Symmetry orbitals

B_{3g} transforms like xy , therefore the two orbitals should have opposite phase:

$$\begin{aligned}\theta_a &= \frac{1}{\sqrt{2}}(\phi_2 - \phi_5) \\ \theta_b &= \frac{1}{\sqrt{2}}(\phi_1 - \phi_6)\end{aligned}\tag{3.1.51}$$

B_{1u} transforms as z , so all orbitals should have the same phase:

$$\begin{aligned}\theta_c &= \frac{1}{\sqrt{2}}(\phi_2 + \phi_5) \\ \theta_d &= \frac{1}{\sqrt{2}}(\phi_1 + \phi_6) \\ \theta_e &= \frac{1}{\sqrt{2}}(\phi_3 + \phi_4)\end{aligned}\tag{3.1.52}$$

B_{2g} transforms like xz , so opposite phase:

$$\theta_f = \frac{1}{\sqrt{2}}(\phi_3 - \phi_4)\tag{3.1.53}$$

Secular equation

The secular equation for B_{3g} symmetry orbitals is

$$\begin{pmatrix} H_{aa} - E & H_{ab} \\ H_{ba} & H_{bb} - E \end{pmatrix} \begin{pmatrix} c_a \\ c_b \end{pmatrix} = 0\tag{3.1.54}$$

We can evaluate the matrix elements H_{ij} very rapidly with the Hückel approximations:

$$H_{aa} = \alpha, \quad H_{ab} = H_{ba} = \beta, \quad H_{bb} = \alpha\tag{3.1.55}$$

So we have

$$\begin{pmatrix} \alpha - E & \beta \\ \beta & \alpha - E \end{pmatrix} \begin{pmatrix} c_a \\ c_b \end{pmatrix} = 0\tag{3.1.56}$$

Which means the secular determinant is zero. A common simplification is to divide throughout by β and setting $(\alpha - E)/\beta = x$:

$$\begin{vmatrix} x & 1 \\ 1 & x \end{vmatrix} = 0\tag{3.1.57}$$

So $E = \alpha - x\beta$

$$E_1 = \alpha + \beta \text{ or } E_2 = \alpha - \beta \quad (3.1.58)$$

The coefficients can be obtained easily so the overall solutions are

$$\begin{aligned} \psi_1 &= \frac{1}{2}(\phi_1 + \phi_2 + \phi_5 + \phi_6) \\ \psi_2 &= \frac{1}{2}(\phi_1 - \phi_2 - \phi_5 + \phi_6) \end{aligned} \quad (3.1.59)$$

The B_{1u} equation is

$$\begin{pmatrix} H_{cc} - E & H_{cd} & H_{ce} \\ H_{dc} & H_{dd} - E & H_{de} \\ H_{ec} & H_{ed} & H_{ee} - E \end{pmatrix} \begin{pmatrix} c_c \\ c_d \\ c_e \end{pmatrix} = 0 \quad (3.1.60)$$

The diagonal matrix elements H_{ii} always come to α , the off-diagonal terms are

$$H_{cd} = \beta \quad H_{ce} = 2\beta \quad H_{de} = 0 \quad (3.1.61)$$

So the simplified secular equation is

$$\begin{pmatrix} x & 1 & 2 \\ 1 & x & 0 \\ 2 & 0 & x \end{pmatrix} \begin{pmatrix} c_c \\ c_d \\ c_e \end{pmatrix} = 0 \quad (3.1.62)$$

$E = \alpha - x\beta$ so $E = \alpha$, $E = \alpha \pm \sqrt{5}\beta$. The coefficients can be solved but is too laborious to get into.

The B_{2g} secular matrix is 1 by 1, so the energy is just $E = \alpha$. We can draw up the energy level diagram as

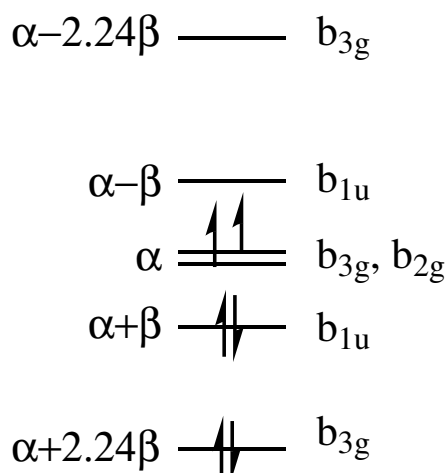


Figure 3.2: There are a total of 6 p_π electrons in the system, with atoms 3 and 4 formally radicals.

And we define the *delocalisation energy* as

Definition 3.1.4.1 (Delocalisation energy). The delocalisation energy is the difference between the energy of the fully delocalised system and the putative system where only localised double bonds exist.

So for this system, as we have seen, a localised double bond MO between the same element have two levels, $E = \alpha + \beta$ and $E = \alpha - \beta$, so a two-electron bond have $2\beta + 2\beta$. In the localised version of this system, there are two two-electron double bonds and two radical atoms, which just have energies of α , so the delocalisation energy is

$$E_{\text{delocalisation}} = E_{\text{delocalised}} - E_{\text{localised}} = 6\alpha + 2(\sqrt{5} + 1)\beta - (6\alpha + 4\beta) \approx 2.47\beta \quad (3.1.63)$$

3.2 PERTURBATION THEORY

3.2.1 Nondegenerate theory

Say we are unable to solve

$$H\psi = E\psi, \quad (3.2.1)$$

however we do have exact solutions for

$$H^{(0)}\psi_n^{(0)} = E_n^{(0)}\psi_n^{(0)}, \quad (3.2.2)$$

where $H^{(0)}$ should be broadly functionally similar to H . We omit the subscript n for now for simplicity's sake. This means we can write the original Hamiltonian as $H^{(0)}$ and the difference:

$$H = H^{(0)} + \lambda H', \quad (3.2.3)$$

where the first term is called the **unperturbed Hamiltonian** and the difference is called the **perturbation**. λ is simply a dimensionless parameter running from 0 to 1, and it should be small for the perturbed system to resemble the unperturbed system. We therefore hope to solve

$$(H^{(0)} + \lambda H')\psi = E\psi. \quad (3.2.4)$$

If λ is small, we can write

$$\psi = \psi^{(0)} + \lambda\psi^{(1)} + \lambda^2\psi^{(2)} + \dots \quad (3.2.5a)$$

$$E = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \dots, \quad (3.2.5b)$$

where, formally (added for clarity, rarely used)

$$E^{(k)} = \frac{1}{k!} \frac{d^k E^{(0)}}{d\lambda^k} \Big|_{\lambda=0} \quad (3.2.6a)$$

$$\psi^{(k)} = \frac{1}{k!} \frac{d^k \psi^{(0)}}{d\lambda^k} \Big|_{\lambda=0}. \quad (3.2.6b)$$

So, the Schrödinger equation has now become

$$\begin{aligned} & (H^{(0)} + \lambda H')(\psi^{(0)} + \lambda\psi^{(1)} + \lambda^2\psi^{(2)} + \dots) \\ &= (E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \dots)(\psi^{(0)} + \lambda\psi^{(1)} + \lambda^2\psi^{(2)} + \dots). \end{aligned} \quad (3.2.7)$$

Expanding and collecting like terms of λ we have

$$H^{(0)}\psi^{(0)} = E^{(0)}\psi^{(0)} \quad (\lambda^0 \text{ terms}) \quad (3.2.8a)$$

$$H^{(0)}\psi^{(1)} + H'\psi^{(0)} = E^{(0)}\psi^{(1)} + E^{(1)}\psi^{(0)} \quad (\lambda^1 \text{ terms}) \quad (3.2.8b)$$

...

Theorem 3.2.1.1 reads, in bra-ket notation

$$H^{(0)}|\psi^{(1)}\rangle + H'|\psi^{(0)}\rangle = E^{(0)}|\psi^{(1)}\rangle + E^{(1)}|\psi^{(0)}\rangle. \quad (3.2.9)$$

Multiplying with the bra $\langle\psi^{(0)}|$ we have

$$\begin{aligned} \langle\psi^{(0)}|H^{(0)}\psi^{(1)}\rangle + \langle\psi^{(0)}|H'\psi^{(0)}\rangle &= E^{(0)}\langle\psi^{(0)}|\psi^{(1)}\rangle + E^{(1)}\langle\psi^{(0)}|\psi^{(0)}\rangle \\ \langle H^{(0)}\psi^{(0)}|\psi^{(1)}\rangle + \langle\psi^{(0)}|H'\psi^{(0)}\rangle &= E^{(0)}\langle\psi^{(0)}|\psi^{(1)}\rangle + E^{(1)} \\ E^{(1)} &= \langle\psi^{(0)}|H'\psi^{(0)}\rangle, \end{aligned} \quad (3.2.10)$$

where in the second line we have use the hermiticity of the Hamiltonian. We now have the theorem

Theorem 3.2.1.1 (First-order perturbation theory). For known perturbed operator H' and known unperturbed eigenfunction $\psi^{(0)}$, the first-order correction in energy, $E^{(1)}$ is given by

$$E^{(1)} = \langle\psi^{(0)}|H'\psi^{(0)}\rangle. \quad (3.2.11)$$

We now have the energy correction, what about the correction in wavefunction? We rewrite Equation 3.2.9 as

$$(H^{(0)} - E_n^{(0)})\psi_n^{(1)} = -(H' - E_n^{(1)})\psi_n^{(0)}, \quad (3.2.12)$$

where we have added the n subscript where it did not really matter before. This is essentially an inhomogeneous differential equation for $\psi_n^{(1)}$. We try expanding $\psi_n^{(1)}$ in $\psi_m^{(0)}$'s as we know they form a complete set:

$$\psi_n^{(1)} = \sum_{m \neq n} c_m^{(n)} \psi_m^{(0)}. \quad (3.2.13)$$

We need to exclude the $m = n$, i.e., $\psi_n^{(0)}$ term because the normalisation requirement for ψ_n requires that

$$1 = \langle\psi_n|\psi_n\rangle = \langle\psi_n^{(0)}|\psi_n^{(0)}\rangle + \lambda(\langle\psi_n^{(1)}|\psi_n^{(0)}\rangle + \langle\psi_n^{(0)}|\psi_n^{(1)}\rangle) + \lambda^2(\dots) + \dots, \quad (3.2.14)$$

but since the first term of the sum must be 1 due to orthonormality of the unperturbed state, the subsequent terms must be zero, which means

$$\langle \psi_n^{(1)} | \psi_n^{(0)} \rangle = 0, \quad (3.2.15)$$

which means $\psi_n^{(1)}$ cannot contain a $\psi_n^{(0)}$ term. Now putting this into Equation 3.2.12 we get

$$\begin{aligned} \sum_{m \neq n} (E_m^{(0)} - E_n^{(0)}) c_m^{(n)} \psi_m^{(0)} &= -(H' - E_n^{(1)}) \psi_n^{(0)} \\ \sum_{m \neq n} (E_m^{(0)} - E_n^{(0)}) c_m^{(n)} \langle \psi_l^{(0)} | \psi_m^{(0)} \rangle &= -\langle \psi_l^{(0)} | H' \psi_n^{(0)} \rangle + E_n^{(1)} \langle \psi_l^{(0)} | \psi_n^{(0)} \rangle \end{aligned} \quad (3.2.16)$$

If we set $l = n$ we recover the energy correction since the left side will return zero. But if $l \neq n$, we get

$$(E_l^{(0)} - E_n^{(0)}) c_l^{(n)} = -\langle \psi_l^{(0)} | H' \psi_n^{(0)} \rangle, \quad (3.2.17)$$

which gives us expression for the $c_m^{(n)}$'s (by relabelling) and so we have what we're looking for:

Theorem 3.2.1.2 (First-order correction in wavefunction).

$$\psi_n^{(1)} = \sum_{m \neq n} \frac{\langle \psi_m^{(0)} | H' \psi_n^{(0)} \rangle}{(E_n^{(0)} - E_m^{(0)})} \psi_m^{(0)}. \quad (3.2.18)$$

Going one power higher we have

$$\begin{aligned} H^{(0)} \psi_n^{(2)} + H' \psi_n^{(1)} &= E_n^{(0)} \psi_n^{(2)} + E_n^{(1)} \psi_n^{(1)} + E_n^{(2)} \psi_n^{(0)} \\ \langle \psi_n^{(0)} | H^{(0)} \psi_n^{(2)} \rangle + \langle \psi_n^{(0)} | H' \psi_n^{(1)} \rangle &= E_n^{(0)} \langle \psi_n^{(0)} | \psi_n^{(2)} \rangle + E_n^{(1)} \langle \psi_n^{(0)} | \psi_n^{(1)} \rangle + E_n^{(2)} \langle \psi_n^{(0)} | \psi_n^{(0)} \rangle \\ E_n^{(0)} \langle \psi_n^{(0)} | \psi_n^{(2)} \rangle + \langle \psi_n^{(0)} | H' \psi_n^{(1)} \rangle &= E_n^{(0)} \langle \psi_n^{(0)} | \psi_n^{(2)} \rangle + E_n^{(1)} \langle \psi_n^{(0)} | \psi_n^{(1)} \rangle + E_n^{(2)} \\ E_n^{(2)} &= \langle \psi_n^{(0)} | H' \psi_n^{(1)} \rangle - E_n^{(1)} \langle \psi_n^{(0)} | \psi_n^{(1)} \rangle \end{aligned} \quad (3.2.19)$$

but as we have already seen in Equation 3.2.15, the second term shall be zero, so

$$\begin{aligned} E_n^{(2)} &= \sum_{m \neq n} c_m^{(n)} \langle \psi_n^{(0)} | H' \psi_m^{(0)} \rangle \\ &= \sum_{m \neq n} \frac{\langle \psi_m^{(0)} | H' \psi_n^{(0)} \rangle \langle \psi_n^{(0)} | H' \psi_m^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \\ &= \sum_{m \neq n} \frac{|\langle \psi_m^{(0)} | H' \psi_n^{(0)} \rangle|^2}{E_n^{(0)} - E_m^{(0)}} \end{aligned} \quad (3.2.20)$$

Theorem 3.2.1.3 (Second-order energy correction).

$$E_n^{(2)} = \sum_{m \neq n} \frac{|\langle \psi_m^{(0)} | H' \psi_n^{(0)} \rangle|^2}{E_n^{(0)} - E_m^{(0)}}. \quad (3.2.21)$$

Now, all of this will crumble if $E_n^{(0)} = E_m^{(0)}$, *i.e.*, if there are two **degenerate** unperturbed states. We will discuss it after we look at some examples.

3.2.2 Examples

Anharmonic oscillators

The Hamiltonian is now given by

$$H = \underbrace{-\frac{\hbar^2}{2\mu} \frac{d^2}{dx^2} + \frac{1}{2} kx^2}_{H^{(0)}} + \underbrace{\frac{1}{6} \gamma x^3 + \frac{1}{24} bx^4}_{H'}. \quad (3.2.22)$$

And we know that the unperturbed ground wavefunction is

$$\psi^{(0)}(x) = \left(\frac{\alpha}{\pi}\right)^{1/4} e^{-\alpha x^2/2} \quad (3.2.23)$$

where $\alpha = (k\mu/\hbar^2)^{1/2}$.

So we apply the first order perturbation theory:

$$\begin{aligned} E^{(1)} &= \langle \psi^{(0)} | H' \psi^{(0)} \rangle \\ &= \left(\frac{\alpha}{\pi}\right)^{1/2} \left[\underbrace{\frac{\gamma}{6} \int_{-\infty}^{+\infty} x^3 e^{-\alpha x^2} dx}_{\text{odd integrand, 0}} + \frac{b}{24} \int_{-\infty}^{+\infty} x^4 e^{-\alpha x^2} dx \right] \\ &= \frac{b}{32\alpha^2} = \frac{\hbar^2 b}{32k\mu}, \end{aligned} \quad (3.2.24)$$

where we have used the standard integral. So the first order ground-state energy is

$$E = \frac{1}{2} \hbar \omega + \frac{1}{32} \frac{\hbar^2 b}{k\mu}. \quad (3.2.25)$$

Helium ground state (again)

The unperturbed Hamiltonian is

$$H^{(0)} = H_1 + H_2 \quad (3.2.26)$$

where the H_i 's are $Z = 2$ hydrogenlike Hamiltonians, and the unperturbed ground state and energy are

$$\begin{aligned} \psi^{(0)} &= \psi_{1s}(r_1, \theta_1, \phi_1) \psi_{1s}(r_2, \theta_2, \phi_2) \\ E^{(0)} &= -\frac{Z^2}{2} - \frac{Z^2}{2}, \end{aligned} \quad (3.2.27)$$

and the perturbed Hamiltonian is

$$H' = \frac{e^2}{4\pi\epsilon r_{12}}. \quad (3.2.28)$$

Applying the theory we have

$$E^{(1)} = \iint \psi_{1s}(\mathbf{r}_1) \psi_{1s}(\mathbf{r}_2) \frac{e^2}{4\pi\epsilon r_{12}} \psi_{1s}(\mathbf{r}_1) \psi_{1s}(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 = \frac{5}{8}Z, \quad (3.2.29)$$

where we have used the same result from Equation 3.1.32. So the first-order energy is

$$E = -Z^2 + \frac{5}{8}Z = -2.75. \quad (3.2.30)$$

Slightly worse than the variational calculation we did earlier, but with higher order corrections the energy will become highly accurate.

3.2.3 Degenerate theory

This section draws heavily from L2.3-3.2 of [4] in conjunction with its accompanying notes ([3]) and also Section 6.8 of [2].

Now suppose we have an N -fold degeneracy of the *unperturbed* wavefunction, which is to say, for $1 < k < N$

$$H^{(0)}|n^{(0)}, k\rangle = E^{(0)}|0, k\rangle, \quad (3.2.31)$$

this is the same as saying that $H^{(0)}$ is, in the basis of the eigenstate, a diagonal matrix with diagonal entries

$$H^{(0)} = \text{diag}\{E_1^{(0)}, E_2^{(0)}, \dots, \underbrace{E_n^{(0)}, \dots, E_n^{(0)}}_N, \dots\}. \quad (3.2.32)$$

In the *degenerate subspace*, \mathbb{V}_N , we choose a collection of N *orthonormal* eigenstates

$$|n^{(0)}; 1\rangle, |n^{(0)}; 2\rangle, \dots, |n^{(0)}; N\rangle. \quad (3.2.33)$$

The other states outside the subspace may also contain degeneracy but we do not have to be concerned with them, since they have *different energy* than *these* degenerate states, and so do not pose problems in the denominator, so to say. We label this space \hat{V} , with eigenvectors (and since $H^{(0)}$ is diagonal in \hat{V} , the basis vectors) $|p^{(0)}\rangle$. Compactly put, the total state space (Hilbert space) is

$$\mathcal{H} = \mathbb{V}_N \oplus \hat{V}. \quad (3.2.34)$$

We will use the following orthonormality conditions:

Property 3.2.3.1 (Orthonormality conditions of DPT).

$$\langle p^{(0)} | q^{(0)} \rangle = \delta_{pq}, \quad \langle p^{(0)} | n^{(0)}; k \rangle = 0, \quad \langle n^{(0)}; k | n^{(p>0)}; k \rangle = 0, \quad (3.2.35)$$

where the third is due to Equation 3.2.15 again, which intuitively says that the higher order terms receive no correction from the ground state, because it can just be reabsorbed into the normalisation of the ground state (just a restatement of the justification of Equation 3.2.15). However note that it can receive correction from *other* degenerate states, *i.e.*, $\langle n^{(0)}; k | n^{(p>0)}; k \rangle$ is not necessarily zero.

In the most general terms, when the perturbation (λ) is turned on, the degenerate states and their energies will become

$$\begin{aligned} |n^{(0)}; k\rangle &\rightarrow |n; k\rangle_\lambda \equiv |n^{(0)}; k\rangle + \lambda |n^{(1)}; k\rangle + \lambda^2 |n^{(2)}; k\rangle + O(\lambda^3) \\ E_n^{(0)} &\rightarrow E_{n,k}(\lambda) \equiv E_n^{(0)} + \lambda E_{n,k}^{(1)} + \lambda^2 E_{n,k}^{(2)} + O(\lambda^3) \end{aligned} \quad (3.2.36)$$

Now, we say that the degeneracy is *lifted* in the first order if all $E_{n,k}^{(1)}$ has different values, *i.e.*, they have now diverged and can be told apart. Our goal is find these corrections and then the first-order corrections in the eigenvector, $|n^{(1)}; k\rangle$, for all k .

To cast this in the eigenfunction language we've used for the nondegenerate case, this is to say that we will determine the linear combinations of $\psi_k^{(0)}$'s that make up $\psi_k^{(1)}$, and implicitly in this we also need to find an *appropriate set* of $\psi_k^{(0)}$'s such that our final linear combinations will be as simple as possible.

The Schrödinger equation to solve is

$$H(\lambda)|n; k\rangle_\lambda = E_{n,k}(\lambda)|n; k\rangle_\lambda, \quad (3.2.37)$$

where

$$H(\lambda) = H^{(0)} + \lambda H', \quad (3.2.38)$$

is just our Hamiltonian with perturbation as before. We'll extract our equations in ascending orders of λ exactly as before:

$$\lambda^0 : (H^{(0)} - E_n^{(0)})|n^{(0)}; k\rangle = 0, \quad (3.2.39a)$$

$$\lambda^1 : (H^{(0)} - E_n^{(0)})|n^{(1)}; k\rangle = (E_{n,k}^{(1)} - H')|n^{(0)}; k\rangle, \quad (3.2.39b)$$

$$\lambda^2 : (H^{(0)} - E_n^{(0)})|n^{(2)}; k\rangle = (E_{n,k}^{(1)} - H')|n^{(1)}; k\rangle + E_{n,k}^{(2)}|n^{(0)}; k\rangle, \quad (3.2.39c)$$

...

The λ^0 equation is trivial, and we can ignore it. Now we have a three-step plan to solve our problem at hand:

1. Right-multiply the λ^1 equation with $\langle n^{(0)}; l|$, conclude condition that the choice of basis must fulfill.
2. Use the λ^1 equation to get the components of $|n^{(1)}; k\rangle$ in \hat{V} .
3. Right-multiply the λ^2 equation with $\langle n^{(0)}; l|$ to get the second order energy correction $E_{n,k}^{(2)}$ and the component of $\langle n^{(0)}; l|$ in \mathbb{V}_N .

Step 1

The LHS is simply zero due to the orthonormality condition and the hermiticity of $H^{(0)}$.

The RHS yields

$$\langle n^{(0)}; l|(E_{n,k}^{(1)} - H')|n^{(0)}; k\rangle = 0 \Rightarrow (H')_{lk} = E_{n,k}^{(1)}\delta_{lk} \quad (3.2.40)$$

Now this equation tells us several very important things, the most important being

Theorem 3.2.3.1 (Diagonalisation of H'). H' is diagonalised in \mathbb{V}_N , and the first order correction in energy is

$$E_n^{(1)} = (H')_{nk,nk}, \quad (3.2.41)$$

the diagonal entries of H' in the choice of basis in \mathbb{V}_N that diagonalises H' . The n subscript is just to remind us that there's an n -fold degeneracy.

There are three additional remarks:

1. We have only shown that H' is diagonalised in \mathbb{V}_N *only*, it is **not** diagonalised on the entire Hilbert space \mathcal{H} , *i.e.*, the basis vectors $|n^{(0)}; k\rangle$ are eigenvectors only in \mathbb{V}_N not \mathcal{H} . We can see this by inserting a resolution of identity:

$$\begin{aligned} H'|n^{(0)}; l\rangle \Big|_{\mathcal{H}} &= \sum_q |n^{(0)}; q\rangle \langle n^{(0)}; q| H'|n^{(0)}; l\rangle \Big|_{\mathbb{V}_N} + \sum_p |p^{(0)}\rangle \langle p^{(0)}| H'|n^{(0)}; l\rangle \Big|_{\hat{V}} \\ &= \sum_q E_{n,l}^{(1)} \delta_{lq} |n^{(0)}; q\rangle + \sum_p |p^{(0)}\rangle \langle p^{(0)}| H'|n^{(0)}; l\rangle \\ &= E_{n,l}^{(1)} |n^{(0)}; l\rangle + \sum_p |p^{(0)}\rangle \langle p^{(0)}| H'|n^{(0)}; l\rangle \end{aligned} \quad (3.2.42)$$

If we only resolved the identity in \mathbb{V}_N we will have gotten the first term only, which means that it's diagonalised in it. The second term means that $|n^{(0)}; k\rangle$ indeed have components along \hat{V} , *i.e.*, it receives contribution from other states, which intuitively makes sense.

2. As we can see, we can *only* determine the energy correction if H' is diagonalised. This requires the choice of a 'good' basis set of $|n^{(0)}; k\rangle$'s, as alluded to before. If we fail to choose a good basis set we just have to manually diagonalise it.
3. This relates to the above point, we can utilise a theorem, introduced below to determine if H' is diagonal without much computation. If it's not, however, we'll have to manually diagonalise it, which will be discussed below.

Theorem 3.2.3.2 (Quick determination of whether H' diagonalises). The matrix H' is diagonal for a choice of basis in \mathbb{V}_N if there is a Hermitian operator K that commutes with H' for which the chosen basis vectors are eigenstates of K with *different* eigenvalues.

Proof. Let $[H', K] = 0$ and that for two basis states in \mathbb{V}_N , $|n^{(0)}; p\rangle$ and $|n^{(0)}; q\rangle$ we have eigenvalues of K , $\lambda_p \neq \lambda_q$, so we have

$$0 = \langle n^{(0)}; p|[H', K]|n^{(0)}; q\rangle = \underbrace{(\lambda_p - \lambda_q)}_{\neq 0} \langle n^{(0)}; p|H'|n^{(0)}; q\rangle. \quad (3.2.43)$$

Therefore the off-diagonal entries of H' must vanish. □

Now, if we weren't so lucky to choose an orthonormal basis, all we need to do is to examine the elements of H' in \mathbb{V}_N and calculate the roots of the *characteristic equation*, which we can

then use to diagonalise the matrix.

Step 2

We now determine the components of $|n^{(1)}; k\rangle$ along \hat{V} , whose basis vectors are $|p^{(0)}\rangle$, so we right-multiply it to the equation to find

$$\begin{aligned} \langle p^{(0)} | (H^{(0)} - E_n^{(0)}) | n^{(1)}; k \rangle \Big|_{\hat{V}} &= -\langle p^{(0)} | H' | n^{(0)}; k \rangle \\ (E_p^{(0)} - E_n^{(0)}) \langle p^{(0)} | n^{(1)}; k \rangle \Big|_{\hat{V}} &= -(H')_{p,nk} \\ |n^{(1)}; k\rangle \Big|_{\hat{V}} &= -\sum_p \frac{(H')_{p,nk}}{E_p^{(0)} - E_n^{(0)}} |p^{(0)}\rangle \\ |n^{(1)}; k\rangle &= |n^{(1)}; k\rangle \Big|_{\mathbb{V}_N} + |n^{(1)}; k\rangle \Big|_{\hat{V}} = -\sum_p \frac{(H')_{p,nk}}{E_p^{(0)} - E_n^{(0)}} |p^{(0)}\rangle + |n^{(1)}; k\rangle \Big|_{\mathbb{V}_N} \end{aligned} \quad (3.2.44)$$

Step 3

Now we want to find the components in the degenerate space, so we right-multiply the λ^3 equation with the basis vectors $\langle n^{(0)}; l|$:

$$\begin{aligned} 0 &= \langle n^{(0)}; l | (E_{nk}^{(1)} - H') | n^{(1)}; k \rangle + E_{nk}^{(2)} \delta_{kl} \\ &= -\langle n^{(0)}; l | (E_{nk}^{(1)} - H') \sum_p |p^{(0)}\rangle \frac{(H')_{p,nk}}{E_p^{(0)} - E_n^{(0)}} \\ &\quad + \langle n^{(0)}; l | (E_{nk}^{(1)} - H') | n^{(1)}; k \rangle \Big|_{\mathbb{V}_N} + E_{nk}^{(2)} \delta_{kl} \\ &= \langle n^{(0)}; l | H' \sum_p |p^{(0)}\rangle \frac{(H')_{p,nk}}{E_p^{(0)} - E_n^{(0)}} + \langle n^{(0)}; l | (E_{nk}^{(1)} - H') | n^{(1)}; k \rangle \Big|_{\mathbb{V}_N} + E_{nk}^{(2)} \delta_{kl} \\ &= \sum_p \frac{(H')_{nl,p} (H')_{p,nk}}{E_p^{(0)} - E_n^{(0)}} + \langle n^{(0)}; l | (E_{nk}^{(1)} - H') | n^{(1)}; k \rangle \Big|_{\mathbb{V}_N} + E_{nk}^{(2)} \delta_{kl}. \end{aligned} \quad (3.2.45)$$

To simplify this expression further we insert a resolution of identity in the second term

$$\begin{aligned} &\langle n^{(0)}; l | H' | n^{(1)}; k \rangle \Big|_{\mathbb{V}_N} \\ &= \underbrace{\langle n^{(0)}; l | H' | \left(\sum_q |n^{(0)}; q\rangle \langle n^{(0)}; q| + \sum_p |p^{(0)}\rangle \langle p^{(0)}| \right)}_{(H')_{nl,nq}} \underbrace{|n^{(1)}; k\rangle \Big|_{\mathbb{V}_N}}_{\hat{V} \perp \mathbb{V}_N \Rightarrow 0} \\ &= \sum_q E_{nl}^{(1)} \delta_{lq} \langle n^{(0)}; q | n^{(1)}; k \rangle \Big|_{\mathbb{V}_N}. \end{aligned} \quad (3.2.46)$$

So we obtain the ‘master’ equation:

$$\underbrace{-\sum_p \frac{(H')_{nl,p} (H')_{p,nk}}{E_p^{(0)} - E_n^{(0)}}}_{\text{known quantities}} + \underbrace{(E_{nk}^{(1)} - E_{nl}^{(1)}) \langle n^{(0)}; l | n^{(1)}; k \rangle \Big|_{\mathbb{V}_N}}_{\text{main unknown}} + E_{nk}^{(2)} \delta_{lk} = 0. \quad (3.2.47)$$

This equation is quite rich in information: we can obtain the second-order energy correction if we set $l = k$, which eliminates our main unknown due to the third orthonormality condition. When we set $l \neq k$ the second-order energy disappears, leaving us a simpler expression for our main unknown.

Theorem 3.2.3.3 (Second-order energy correction, degenerate). When $l = k$, we immediately obtain

$$\begin{aligned} E_{nk}^{(2)} &= \langle n^{(0)}; k | H' | n^{(1)}; k \rangle \Big|_{\hat{V}} \\ &= - \sum_p \frac{|(H')_{nk,p}|^2}{E_p^{(0)} - E_n^{(0)}}, \end{aligned} \quad (3.2.48)$$

We see that it's the same form as the nondegenerate case (Theorem 3.2.1.3), only that the sum is over outside the degenerate space.

As an aside, if we believed naively that there are no components inside the degenerate space, we would have concluded that

$$\langle n^{(0)}; l | H' | n^{(1)}; k \rangle \Big|_{\hat{V}} = 0, \quad (3.2.49)$$

which we have shown to be not zero.

Finally, we set $l \neq k$ to get

$$\begin{aligned} \langle n^{(0)}; l | n^{(1)}; k \rangle \Big|_{\mathbb{V}_N} &= \frac{1}{E_{nk}^{(1)} - E_{nl}^{(1)}} \langle n^{(0)}; l | H' | n^{(1)}; k \rangle \Big|_{\hat{V}} \\ &= - \frac{1}{E_{nk}^{(1)} - E_{nl}^{(1)}} \sum_p \frac{(H')_{nl,p} (H')_{p,nk}}{E_p^{(0)} - E_n^{(0)}}, \quad l \neq k \end{aligned} \quad (3.2.50)$$

The fraction implies that the degeneracy must be *completely* lifted, otherwise we must resolve the degeneracy to higher orders. Projecting along the basis we have

$$|n^{(1)}; k\rangle \Big|_{\mathbb{V}_N} = - \sum_{l \neq k} |n^{(0)}; k\rangle \frac{1}{E_{nk}^{(1)} - E_{nl}^{(1)}} \sum_p \frac{(H')_{nl,p} (H')_{p,nk}}{E_p^{(0)} - E_n^{(0)}}. \quad (3.2.51)$$

Curiously the degenerate corrections depend on the nondegenerate corrections. In all,

Theorem 3.2.3.4 (First-order correction to eigenfunction, degenerate).

$$\begin{aligned} |n^{(1)}; k\rangle &= |n^{(1)}; k\rangle \Big|_{\mathbb{V}_N} + |n^{(1)}; k\rangle \Big|_{\hat{V}} \\ &= - \sum_{l \neq k} |n^{(0)}; k\rangle \frac{1}{E_{nk}^{(1)} - E_{nl}^{(1)}} \sum_p \frac{(H')_{nl,p} (H')_{p,nk}}{E_p^{(0)} - E_n^{(0)}} - \sum_p \frac{(H')_{p,nk}}{E_p^{(0)} - E_n^{(0)}} |p^{(0)}\rangle \end{aligned} \quad (3.2.52)$$

3.3 THE HARTREE-FOCK METHOD

3.3.1 Self-consistent field method

A more inspired trial function for the variational method could be the **Slater orbitals**, given by

$$S_{nlm}(r, \theta, \phi) = N r^{n-1} e^{-\zeta r} Y_l^m(\theta, \phi), \quad (3.3.1)$$

where ζ is a variational parameter and

$$N = \frac{(2\zeta)^{n+1/2}}{\sqrt{(2n)!}}. \quad (3.3.2)$$

There is no inclusion of the Legendre polynomial so there aren't any radial nodes, and as a consequence these orbitals aren't in general orthogonal.

Now the starting point of the Hartree-Fock procedure for helium is to write the two-electron wavefunction as a product of two orbitals, typically linear combinations of Slater orbitals:

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \phi(\mathbf{r}_1)\phi(\mathbf{r}_2). \quad (3.3.3)$$

The two one-electron wavefunctions are the same, which is *in direct violation* of the Pauli exclusion principle. We ignore it for the moment. We must try to account for the inter-electronic repulsion in the Hamiltonian:

$$H_1 = -\frac{1}{2}\nabla_1^2 - \frac{2}{r} + V_1^{\text{eff}}(r_1), \quad (3.3.4)$$

where we have adopted atomic units and where the effective potential is where we introduce interelectronic repulsion as an average potential (this is also known as the *mean field approximation*):

$$V_1^{\text{eff}}(r_1) = \langle \phi(\mathbf{r}_2) | \frac{1}{r_{12}} | \phi(\mathbf{r}_2) \rangle. \quad (3.3.5)$$

Now we have our Schrödinger equation

$$H_1\phi(\mathbf{r}_1) = \epsilon_1\phi(\mathbf{r}_1). \quad (3.3.6)$$

This is called the **Hartree-Fock equation** for a Helium atom. A special feature of the equation is that its Hamiltonian depends on $\phi(\mathbf{r}_2)$ through Equation 3.3.4. $\phi(\mathbf{r}_2)$ in turn is the exact same function as the solution of the equation, $\phi(\mathbf{r}_1)$. This recursive relation allows for iteration and that's the essence of the **self-consistent field method**: we first guess a form, any form, for $\phi(\mathbf{r}_2)$ and evaluate the effective potential, and hence the Hamiltonian,

and we solve the Schrödinger equation for $\phi(\mathbf{r}_1)$, by **todo/todo-supo: check if by variational principle or perturbation theory**, and use the output as the input again, until the output is sufficiently close to the input, or *self-consistent*.

3.3.2 Accounting for spin

In treating the helium atom just now we only used the spatial part of the trial wavefunction, and it appeared to have contradicted Pauli exclusion principle. But we have not, we just need to say that the two electrons must be in the singlet configuration, which is antisymmetric wrt electron exchange. The spin does not contribute anything to any integrals involving spatial integration as the Hamiltonian does not depend on it¹ and that they are made orthonormal.

However, all this triplet and singlet talk is when we only have two electrons, what if we need to construct antisymmetric wavefunctions out of N -electrons? We introduce the

Definition 3.3.2.1 (Slater determinant). We can use the Slater determinant to construct antisymmetric N -electron wavefunctions:

$$\Psi(1, 2, \dots, N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} u_1(1) & u_2(1) & \cdots & u_N(1) \\ u_1(2) & u_2(2) & \cdots & u_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ u_1(N) & u_2(N) & \cdots & u_N(N) \end{vmatrix}, \quad (3.3.7)$$

where $u_i(j)$ are the i -th complete (*i.e.*, spin included) orthonormal orbitals for particle j .

The determinant works because

1. It changes signs whenever two electrons (rows) are swapped;
2. It vanishes whenever two electrons occupy the same state (two identical columns).

For example, we look at the lithium atom. The intuition tells us to try

$$\psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \phi_{1s}(\mathbf{r}_1)\phi_{1s}(\mathbf{r}_2)\phi_{1s}(\mathbf{r}_3), \quad (3.3.8)$$

¹This is again assuming spin and orbitals are uncoupled.

but we the Pauli exclusion principle says we can't possibly accomodate three electrons in the $1s$ orbital, and we put our electron in the next lowest energy orbital², the $2s$, and construct the antisymmetric wavefunction as follows:

$$\psi = \frac{1}{\sqrt{3!}} \begin{vmatrix} 1s\alpha(1) & 1s\beta(1) & 2s\alpha(1) \\ 1s\alpha(2) & 1s\beta(2) & 2s\alpha(2) \\ 1s\alpha(3) & 1s\beta(3) & 2s\alpha(3) \end{vmatrix}, \quad (3.3.9)$$

where $1s$ is shorthand for ψ_{1s} and α is shorthand for spin-up and β for spin-down.

Now let's treat the first excited state of helium, $1s^1 2s^1$.

First we introduce a more compact notation[15]:

$$1s \equiv 1s\alpha, \quad \overline{1s} \equiv 1s\beta \quad (3.3.10)$$

and

$$|1s \overline{2s}\rangle \equiv N \begin{vmatrix} 1s(1) & \overline{2s}(1) \\ 1s(2) & \overline{2s}(2) \end{vmatrix}, \quad (3.3.11)$$

where the normalisation constant is implied. Now let's construct the antisymmetric wavefunctions:

$$|1s 2s\rangle = \frac{1}{\sqrt{2}} \begin{vmatrix} 1s\alpha(1) & 2s\alpha(1) \\ 1s\alpha(2) & 2s\alpha(2) \end{vmatrix} \propto [1s(1)2s(2) - 2s(1)1s(2)](\uparrow\uparrow) \quad ({}^3\Psi, M_s = +1) \quad (3.3.12)$$

$$|1s \overline{2s}\rangle = \frac{1}{\sqrt{2}} \begin{vmatrix} 1s\alpha(1) & 2s\beta(1) \\ 1s\alpha(2) & 2s\beta(2) \end{vmatrix} \propto 1s(1)2s(2) \uparrow\downarrow - 2s(1)1s(2) \downarrow\uparrow \quad (3.3.13)$$

Now this is a state with $M_s = 0$ according to Equation 1.4.228, if we combine a particle with spin up and one with spin down we get zero z -momentum. And here's one more

$$|\overline{1s} 2s\rangle \propto 1s(1)2s(2) \downarrow\uparrow - 2s(1)1s(2) \uparrow\downarrow. \quad (3.3.14)$$

Now, this is the **decoupled picture**, in which m , the z -spin momenta are specified, which is a perfectly legitimate way to specify the system, but we would like to speak of 'singlet' and 'triplet', *i.e.*, of the **coupled picture** where s , the total spin momentum is specified, to

²This is still just to construct a trial wavefunction because we don't know for sure what the energy level is like for lithium atoms, but it's the best guess we've got.

do so we take the linear combination of our states:

$$\begin{aligned}\frac{1}{\sqrt{2}}(|1s \overline{2s}| + |\overline{1s} 2s|) &= [1s(1)2s(2) - 2s(1)1s(2)](\uparrow\downarrow + \downarrow\uparrow) \quad (^3\Psi, M_s = 0) \\ \frac{1}{\sqrt{2}}(|1s \overline{2s}| - |\overline{1s} 2s|) &= [1s(1)2s(2) - 2s(1)1s(2)](\uparrow\downarrow - \downarrow\uparrow) \quad (^1\Psi, M_s = 0)\end{aligned}\quad (3.3.15)$$

The last state is

$$|\overline{1s} \overline{2s}| \propto [1s(1)2s(2) - 2s(1)1s(2)] \downarrow\downarrow \quad (^3\Psi, M_s = -1). \quad (3.3.16)$$

Ok, let's apply first-order perturbation theory to these orbitals. **todo: read atkinsqm on degenerate cases, dk if this is entirely correct, priority: high:**

$$\begin{aligned}E^{(1)} &= \langle \Psi_{\pm} | \frac{1}{r_{12}} | \Psi_{\pm} \rangle \\ &= \frac{1}{2} \langle 1s(1)2s(2) \pm 2s(1)1s(2) | \frac{1}{r_{12}} | 1s(1)2s(2) \pm 2s(1)1s(2) \rangle \\ &\equiv J \pm K,\end{aligned}\quad (3.3.17)$$

where J is the **Coulomb integral** and K the **exchange integral** (Section 7.9 of [2]):

$$\begin{aligned}J &= \langle 1s(1)2s(2) | \frac{1}{r_{12}} | 1s(1)2s(2) \rangle \\ K &= \langle 1s(1)2s(2) | \frac{1}{r_{12}} | 1s(2)2s(1) \rangle\end{aligned}\quad (3.3.18)$$

3.3.3 The Hartree-Fock method

In Section 3.3.1 we have discussed the Hartree-Fock method for the helium atom. However it is only a special case as the spin and spatial wavefunction neatly separates:

$$\Psi = \frac{1}{\sqrt{2}} \begin{vmatrix} 1s\alpha(1) & 1s\beta(1) \\ 1s\alpha(2) & 1s\beta(2) \end{vmatrix} = \frac{1}{\sqrt{2}} 1s(1)1s(2)(\uparrow\downarrow - \downarrow\uparrow). \quad (3.3.19)$$

And it is readily seen that the same is impossible for even lithium atom as the the dependences of all the spatial wavefunctions will be permuted. **todo-suppo: can we introduce LCs like above?**

Now we are ready to introduce the general Hartree-Fock method. Our goal is still the same: to solve the Schrödinger equation with Hamiltonian (same as in Equation 3.0.1):

$$H = \sum_i H^{(0)}(i) + \frac{1}{2} \sum_{i \neq j} \frac{1}{r_{ij}}, \quad (3.3.20)$$

where the first term is the *core Hamiltonian* with nuclear charge Ze for the i -th electron, and the second term is the interelectronic repulsion. By using the Slater determinant to approximate the orbitals and by using the mean-field approximations for effective inter-electronic potentials, and through some tricky derivation (see pp. 528-531 of [2]) the completely general Hartree-Fock equation is introduced as follows:

Theorem 3.3.3.1 (Hartree-Fock equation). The Hartree-Fock equation for a *spin-orbital*^a $\phi_s(1)$ occupied by electron 1 is

$$\left[H^{(0)}(1) + \sum_r (J_r - K_r) \right] \phi_s(1) = \epsilon_s \phi_s(1), \quad (3.3.21)$$

where the sum is over all occupied spin-orbitals.

^aThe spins will produce Kronecker deltas anyway, but it's just more correct to speak of spin-orbitals.

J and K are two operators, defined as follows:

Definition 3.3.3.1 (Coulomb and exchanger operators).

J_r is the **Coulomb operator**

$$J_r |\phi_s(1)\rangle = \langle \phi_r(2) | \frac{1}{r_{12}} | \phi_r(2) \rangle |\phi_s(1)\rangle. \quad (3.3.22)$$

K_r is the **exchange operator**

$$K_r |\phi_s(1)\rangle = \langle \phi_r(2) | \frac{1}{r_{12}} | \phi_s(2) \rangle |\phi_r(1)\rangle. \quad (3.3.23)$$

J accounts for the effect of Coulombic repulsion and K , spin correlation.

Note that since

$$J_a(1) |\phi_a(1)\rangle = K_a(1) |\phi_a(1)\rangle, \quad (3.3.24)$$

the sum in the Fock operator correctly gives results for all electron pair interactions without counting the same electron interacting with itself.

Now the Fock operator depends on n spin-orbitals, and again we solve the Hartree-Fock equation by the self-consistent field method. After that, the Fock operator is theoretically a well-defined Hermitian operator and we can generate an infinite set of eigenfunctions, but in practice we only solve $m \geq n$ optimised spin-orbitals, with the n lowest energy orbitals called **occupied orbitals** and the rest **virtual orbitals**. The Slater determinant (or a linear combination of them, to correct the spin symmetry) composed with the optimised n occupied orbitals is then the HF ground-state wavefunction for the system, Φ_0 .

The HF-SCF method takes into account spin correlation, however, the Coulombic interactions between electrons are treated in an average way, so no instantaneous interactions, nor quantum mechanical electron distribution, are taken into account. These are collectively called **electron correlation**, and is entirely left out of the picture by the HF-SCF method.

Chapter 4

Group theory

4.1 THE BARE MINIMUM

4.1.1 Symmetry elements and symmetry operations

Definition 4.1.1.1 (Symmetry operations). A **symmetry operation** is an operation about a **symmetry element** that leaves the object (molecule) *apparently* unchanged. The actual particles may well have changed places but the important thing is that we can't tell.

Definition 4.1.1.2 (Symmetry elements). A **symmetry element** is said to *generate* symmetry operations, for example, an axis of rotation can generate rotation.

There are only 5 types of symmetry operations and the symmetry elements that generate them:

| Operation | Element |
|--|-------------------------------|
| E , the identity | The object itself |
| C_n , the n-fold rotation | The axis of symmetry |
| σ , the reflection | The mirror plane |
| i , the inversion | The centre of symmetry |
| S_n , the n-fold improper rotation | The axis of improper rotation |

Table 4.1: The 5 symmetry operations and their generating elements.

Some important nomenclature:

1. C_n

- (a) The element with highest n is called the **principle axis**.
- (b) For $n > 2$, an axis can generate two directions of rotation, for example, we have $2C_4$ belonging to the D_{4h} group.

2. σ

- (a) When the element (mirror plane) includes the principle axis, it's called a **vertical plane**, σ_v .
- (b) When the element is perpendicular to the principle axis, it's called a **horizontal plane**, σ_h .
- (c) A special type of vertical planes is when it bisects the angle between two C_2 axes. It's then called a **dihedral plane**, σ_d .

4.1.2 Point groups

We use a symmetry flow chart to determine the point group a molecule belongs to: Example molecules can be found in [23].

4.1.3 Representations

Definition 4.1.3.1 (Matrix representations). A **matrix representation** is a set of **matrix representatives**, which presents the effect of a *symmetry operation* on a molecule, in an **arbitrarily chosen basis**.

The choice of basis, although technically arbitrary, is usually just vectors along the bonds pointing from the central atom or the atomic orbitals, amongst other conventions for other applications.

For any arbitrarily chosen basis, a matrix representation can be generated. We can use the results from the great and little orthogonality theorems (see 5.10-5.11 of [2]) to *reduce* a representation to see how many of each irreducible representation is present in the direct sum. Due to a theorem in group theory that states that the character (trace) of a representative is invariant under a change of basis (a familiarity transform), we can work with characters instead of matrices. An essential theorem on reducing representations (more accurately, characters of representations) is as follows:

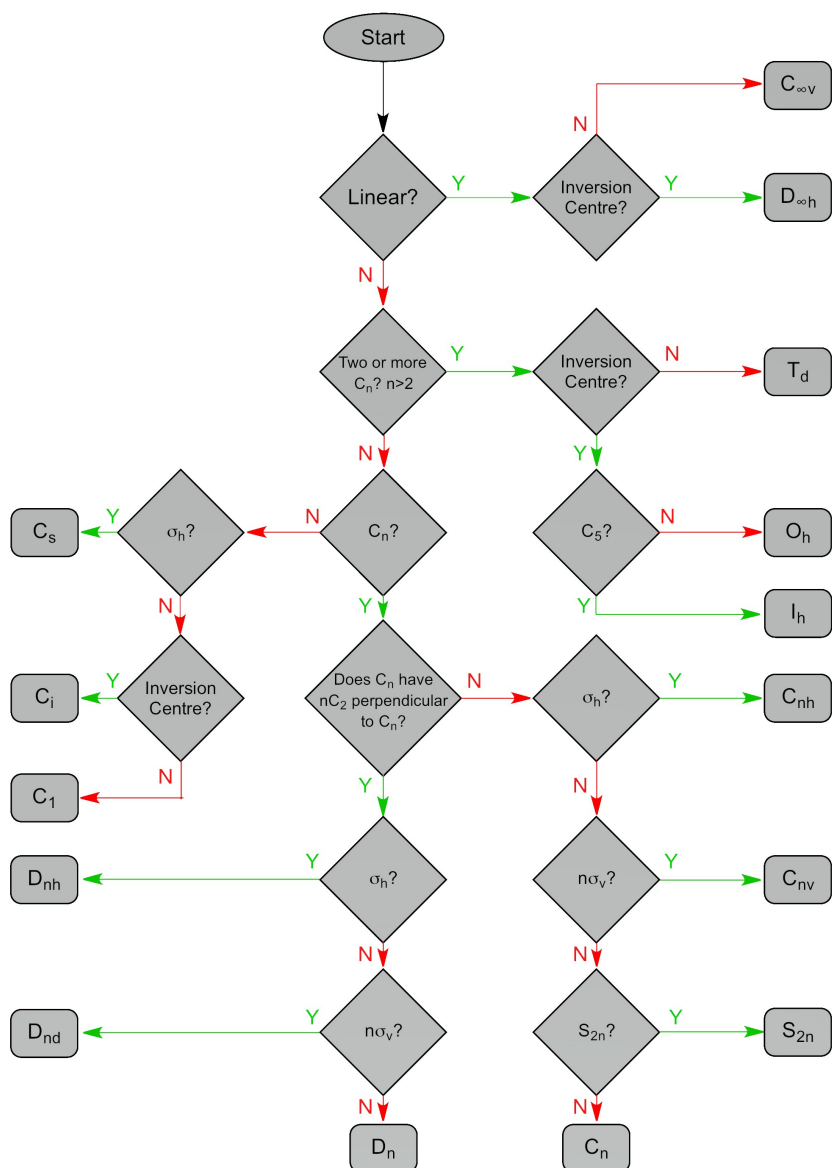


Figure 4.1: Flowchart used to determine point groups.

$$a(I) = \frac{1}{h} \sum_{\text{all classes}} \chi(R)\chi(I)N \quad (4.1.1)$$

N = number of symmetry operations in the class

4.2 MATHEMATICAL FORMULATION

4.2.1 Notations

See the following table for the notation we will employ:

| Meaning | Notation |
|------------------------------|--|
| Basis (a row vector) | $\mathbf{f} = (f_1, f_2, \dots, f_d)$ |
| Symmetry operations | R, S, T, \dots |
| Representation of operations | $\mathbf{D}(R)$ |
| Operation on basis | $R\mathbf{f} \equiv \mathbf{f}\mathbf{D}(R)$ |
| Character of representation | $\chi(R) \equiv \text{tr}\mathbf{D}(R)$ |

Table 4.2: Notations used in group theory.

4.2.2 Basic theorems

We now go on to establish some fundamental lemmas to build on later.

Lemma 4.2.2.1 (Representation of group multiplication). If $RS = T$, then $\mathbf{D}(R)\mathbf{D}(S) = \mathbf{D}(T)$.

Proof.

$$\begin{aligned}
 RS\mathbf{f} &= R[\mathbf{f}\mathbf{D}(S)] = (R\mathbf{f})\mathbf{D}(S) = \mathbf{f}\mathbf{D}(R)\mathbf{D}(S) = T\mathbf{f} = \mathbf{f}\mathbf{D}(T) \\
 &\Rightarrow \mathbf{D}(R)\mathbf{D}(S) = \mathbf{D}(T)
 \end{aligned} \tag{4.2.1}$$

□

Lemma 4.2.2.2 (Representation of the inverse of operation). The representative of the inverse of an operation is the inverse of the representative:

$$\mathbf{D}(R^{-1}) = \mathbf{D}(R)^{-1}. \tag{4.2.2}$$

Proof. Because we know that

$$RR^{-1} = E \tag{4.2.3}$$

so

$$D(R)D(R^{-1}) = D(E) = I \quad (4.2.4)$$

therefore we can identify $D(R^{-1})$ with $D(R)^{-1}$. \square

Oftentimes we would like to change the basis of a representation to simplify it, *i.e.*, to reduce it. We also call a change of basis a **similarity transformation**. An example could be the NH_3 molecule, where we originally choose a basis set of the three s -orbitals. But, as it will turn out, this is not the basis for the irreducible representation, and to find it we need to compute the symmetry-adapted bases of the molecule ([update link](#)). Now say we have the new basis, given by

$$\begin{aligned} s_N &= s_N \\ s_1 &= s_A + s_B + s_C \\ s_2 &= 2s_A - s_B - s_C \\ s_3 &= s_B - s_C \end{aligned}$$

or, in matrix notation

$$f' \equiv f \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 1 & -1 & -1 \end{bmatrix}}_{\equiv c} \quad (4.2.5)$$

the question is then, how to we find the representative, $D(R)$ in this new basis?

Theorem 4.2.2.1 (Similarity transformation). We say that two representations are **similar** if the representatives for the two bases are related by the **similarity transformation**:

$$D(R) = cD'(R)c^{-1} \leftrightarrow D'(R) = c^{-1}D(R)c. \quad (4.2.6)$$

Proof.

$$\begin{aligned} Rf' &= f'D'(R) \\ Rfc &= fcD'(R) \\ Rf &= fcD'(R)c^{-1} = fD(R) \end{aligned} \quad (4.2.7)$$

The result is immediately implied. \square

We introduce a lemma from linear algebra:

Lemma 4.2.2.3 (Trace is invariant under cyclic permutation).

$$\text{tr}ABC = \text{tr}BCA = \text{tr}CAB. \quad (4.2.8)$$

Proof.

$$\text{tr}ABC = (ABC)_{ii} = A_{ik}B_{jk}C_{ki}, \quad (4.2.9)$$

where we have used the Einstein summation convention. Under cyclic permutation, the subscript continue to match, hence giving invariant trace. \square

Theorem 4.2.2.2 (Character is invariant under similarity transform).

$$\chi(R) = \chi'(R). \quad (4.2.10)$$

Proof.

$$\chi(R) = \text{tr}cD'(R)c^{-1} = \text{tr}D'(R) = \chi'(R). \quad (4.2.11)$$

\square

We need to categorise operations into classes, which is defined by

Definition 4.2.2.1 (Classes). Operations R and R' are said to be in the same class, or are *conjugate*, if they are related as follows:

$$R' = S^{-1}RS, \quad (4.2.12)$$

which is to say, they are the same *kind* of operation but performed about symmetry elements that are *related by a symmetry operation*.

Theorem 4.2.2.3 (Operations in the same class have the same character).

$$\chi(R') = \chi(R). \quad (4.2.13)$$

Proof.

$$\chi(R') = \text{tr}D(R') = \text{tr}D^{-1}(S)D(R)D(S) = \text{tr}D(R) = \chi(R). \quad (4.2.14)$$

\square

4.2.3 Irreps and orthogonality theorems

We return to the example of NH_3 , which belongs to the group C_{3v} . In the basis of (s_N, s_1, s_2, s_3) , all the representatives of the group operations all have two diagonal 1's at the top left. The remaining matrices are

$$\begin{array}{ccc} D(E) & D(C_3^+) & D(C_3^-) \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & \begin{bmatrix} -1/2 & -1/2 \\ 1/1 & -1/2 \end{bmatrix} & \begin{bmatrix} -1/2 & 1/2 \\ -1/2 & -1/2 \end{bmatrix} \\ D(\sigma_v) & D(\sigma'_v) & D(\sigma''_v) \\ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} & \begin{bmatrix} -1/2 & 1/2 \\ 3/2 & 1/2 \end{bmatrix} & \begin{bmatrix} -1/2 & -1/2 \\ -3/2 & 1/2 \end{bmatrix} \end{array}$$

We can write this as

$$D^{(4)} = D^{(1)} \oplus D^{(1)} \oplus D^{(2)} \quad (4.2.15)$$

All three of them are **irreducible representations** of group C_{3v} . We know that $D^{(2)}$ is an irreducible representation because the set of two by two matrices are not simultaneously diagonalisable, *i.e.*, they do not commute. [todo: update link to relevant qm section](#)

The basis for the first two $D^{(1)}$'s are s_N and s_1 respectively, we call them

Definition 4.2.3.1 (Unfaithful representation). An **unfaithful representation** of a group is a trivial one by one matrix with 1 as the sole element.

We can see that the characters for them are $(1, 1, 1, 1, 1, 1)$, which means they belong to the A_1 symmetry species. The remaining irreps have characters $(2, -1, -1, 0, 0, 0)$, which put them under the species E . A_1 and E are also called $\Gamma^{(1)}$ and $\Gamma^{(3)}$ respectively.

Now we proceed to introduce the two very important orthogonality theorems:

Theorem 4.2.3.1 (The great orthogonality theorem (GOT)). For a group of order h , let $D^{(\ell)}(R)$ be the representative of the operation R in a d_ℓ -dimensional irrep of symmetry species $\Gamma^{(\ell)}$ of the group, then

$$\sum_R D_{ij}^{(\ell)}(R)^* D_{i'j'}^{(\ell')}(R) = \frac{h}{d_\ell} \delta_{\ell\ell'} \delta_{ii'} \delta_{jj'}. \quad (4.2.16)$$

This is really a postulate at this point because we will not prove it. For most practical

purposes though we just need the

Theorem 4.2.3.2 (The little orthogonality theorem (LOT)).

$$\sum_R \chi^{(\ell)}(R)^* \chi^{(\ell')}(R) = h \delta_{\ell\ell'}. \quad (4.2.17)$$

Proof. We set $j = i$ and $j' = i'$, and sum over diagonal entries since we need the trace:

$$\begin{aligned} \sum_{i,i'} \sum_R D_{ii}^{(\ell)}(R)^* D_{i'i'}^{(\ell')}(R) &= \sum_R \chi^{(\ell)}(R)^* \chi^{(\ell')}(R) \\ &= \sum_{i,i'} \frac{h}{d_\ell} \delta_{\ell\ell'} \delta_{ii'} \delta_{jj'} \\ &= h \delta_{\ell\ell'}, \end{aligned} \quad (4.2.18)$$

we used the fact that the representation is of dimension d_ℓ . □

We can simplify the theorem further by writing

$$\sum_c g(c) \chi^{(\ell)}(c)^* \chi^{(\ell')}(c) = h \delta_{\ell\ell'}, \quad (4.2.19)$$

where c is the class and $g(c)$ is the number of operations in the class. Setting $\ell = \ell'$ we have

$$\sum_c g(c) |\chi^{(\ell)}(c)|^2 = h. \quad (4.2.20)$$

This says that *the sum of squares of the characters of any irreps of a group is equal to the order of the group.*

Recasting this in a vector, we can say that $\{g(c)\}^{1/2} \chi_c^{(\ell)}$ is a component $v_c^{(\ell)}$ of the vector $\mathbf{v}^{(\ell)}$, whose components are distinguished by the subscript c . The LOT can now be written as

$$\mathbf{v}^{(\ell)*} \cdot \mathbf{v}^{(\ell')} = h \delta_{\ell\ell'}. \quad (4.2.21)$$

This is saying that the two vectors are orthogonal unless $\ell' = \ell$. However, the number of orthogonal vectors cannot exceed the dimension of the space (can be far less if some of them are linearly dependent). Another analysis from the GOT tells that the two are actually equal. So we reach

Theorem 4.2.3.3 (Restriction 1). The number of symmetry species is equal to the number of classes.

This analysis can be extended to GOT, with $D_{ij}^{(\ell)}(R)$ being the R -th component of a vector \mathbf{v} identified by three indices ℓ, i and j :

$$\mathbf{v}^{(\ell,i,j)*} \cdot \mathbf{v}^{(\ell',i',j')} = \frac{h}{d_\ell} \delta_{\ell\ell'} \delta_{ii'} \delta_{jj'}. \quad (4.2.22)$$

As the component index is R , we know the vectors are h -dimensional. Also, for a given ℓ (species/irrep), there are d_ℓ^2 vectors. Therefore, a similar argument as the one above concludes that

$$\sum_{\ell} d_\ell^2 \leq h. \quad (4.2.23)$$

We assert that the equality in fact holds, so

Theorem 4.2.3.4 (Restriction 2).

$$\sum_{\ell} d_\ell^2 = h. \quad (4.2.24)$$

4.2.4 Reducing representation

We wish to find a suitable similarity transform that enables us to write a representation as a direct sum of irreps:

$$\mathbf{D}(R) \leftrightarrow \mathbf{D}'(R) = \mathbf{D}^{(\Gamma^{(1)})}(R) \oplus \mathbf{D}^{(\Gamma^{(2)})}(R) \oplus \dots \quad (4.2.25)$$

We introduce the notation

$$\Gamma = \sum_{\ell} a_{\ell} \Gamma^{(\ell)}, \quad (4.2.26)$$

where the sum is really a direct sum and multiplication is repeated direct sums and a_{ℓ} is how many times the irrep appears in the direct sum.

We realise that we need not necessarily find the similarity transform to find the coefficient a_{ℓ} , because we can invoke Theorem 4.2.2.2, and we can say

$$\chi(R) = \chi'(R) = \sum_{\ell} a_{\ell} \chi^{(\ell)}(R). \quad (4.2.27)$$

Now, determining the coefficient is within reach. Let's multiply both sides by $\chi^{(\ell')}(R)^*$

and sum over all elements of the group:

$$\begin{aligned}\sum_R \chi^{(\ell')}(R)^* \chi(R) &= \sum_R \sum_{\ell} a_{\ell} \chi^{(\ell')}(R)^* \chi^{(\ell)}(R) \\ &= h \sum_{\ell} a_{\ell} \delta_{\ell\ell'} = h a_{\ell'}.\end{aligned}\quad (4.2.28)$$

Therefore

Theorem 4.2.4.1 (Reduction of representations).

$$a_{\ell} = \frac{1}{h} \sum_R \chi^{(\ell)}(R)^* \chi(R). \quad (4.2.29)$$

In terms of classes,

$$a_{\ell} = \frac{1}{h} \sum_c g(c) \chi^{(\ell)}(c)^* \chi(c). \quad (4.2.30)$$

4.2.5 Symmetry-adapted bases

After establishing the coefficients we may wish to find out what the similarity transformation was (which was not necessary if we just wanted to know the coefficients), in other words, we can find out the basis of the new representation. The basis is called **symmetry-adapted basis** and the basis functions are called **symmetry-adapted linear combinations** (SALC).

We need to define

Definition 4.2.5.1 (Projection operator). The **projection operator** is defined as

$$P_{ij}^{(\ell)} = \frac{d_{\ell}}{h} \sum_R D_{ij}^{(\ell)}(R)^* R. \quad (4.2.31)$$

Property 4.2.5.1 (Effect of the projection operator).

$$P_{ij}^{(\ell)} f_{j'}^{(\ell')} = f_i^{(\ell)} \delta_{\ell\ell'} \delta_{jj'} \quad (4.2.32)$$

Proof. Consider $\mathbf{f}^{(\ell')} = (f_1^{(\ell')}, f_2^{(\ell')}, \dots, f_d^{(\ell')})$ that form a basis for a $d_{\ell'}$ -dimensional irrep $D^{(\ell')}$ of species $\Gamma^{(\ell')}$ of a group of order h . The effect of any operation of the group is

$$R f_{j'}^{(\ell')} = \sum_{i'} f_{i'}^{(\ell')} D_{i'j'}^{(\ell')}(R). \quad (4.2.33)$$

Now we multiply through by an element of another representative of the same operation, $D_{ij}^{(\ell)}(R)$, and then sum over all elements:

$$\begin{aligned}
 \sum_R D_{ij}^{(\ell)}(R)^* R f_{j'}^{(\ell')} &= \sum_R \sum_{i'} D_{ij}^{(\ell)}(R)^* f_{i'}^{(\ell')} D_{i'j'}^{(\ell')}(R) \\
 &= \sum_{i'} f_{i'}^{(\ell')} \left[\sum_R D_{ij}^{(\ell)}(R)^* D_{i'j'}^{(\ell')}(R) \right] \\
 &= \sum_{i'} f_{i'}^{(\ell')} \left(\frac{h}{d_{\ell'}} \right) \delta_{\ell\ell'} \delta_{ii'} \delta_{jj'} \quad (4.2.34) \\
 &= \left(\frac{h}{d_{\ell'}} \right) \delta_{\ell\ell'} \delta_{jj'} f_i^{(\ell')} \\
 &= \left(\frac{h}{d_{\ell}} \right) \delta_{\ell\ell'} \delta_{jj'} f_i^{(\ell)},
 \end{aligned}$$

where in the third equality we applied GOT and in the last equality we changed ℓ' to ℓ because we might as well due to the presence of $\delta_{\ell\ell'}$. \square

The reason P is called the projection operator because it requires that the function it acts on to be the j -th function of the basis set of $\Gamma^{(\ell)}$, which then converts (projects) it to the i -th basis function, otherwise, it returns zero (orthogonal). This gives us the ability to extract all members of the basis set out of a single member, so long as we know the content of the operator.

A special and important case is when $\ell' = \ell$ and $i = j$, the effect of the operator is then

$$P_{ii}^{(\ell)} f_{j'}^{(\ell)} = f_i^{(\ell)} \delta_{ij'}. \quad (4.2.35)$$

We return to this result very soon.

Now, if we have a set of linearly independent but otherwise arbitrary set of functions $\mathbf{f} = (f_1, f_2, \dots)$, which can serve as the basis of a reducible representation, waiting to be reduced. Just as the new basis functions can be expressed as linear combinations of the original basis functions, we can turn it around and write

$$f_j = \sum_{\ell', j'} f_{j'}^{(\ell')}, \quad (4.2.36)$$

where the sum is also over ℓ' as we're talking about the entire matrix now, also the coefficients are absorbed into the terms. Now let's see the effect of

$$P_{ii}^{(\ell)} f_j = \sum_{\ell', j'} P_{ii}^{(\ell)} f_{j'}^{(\ell')} = \sum_{\ell', j'} \delta_{\ell\ell'} \delta_{ij'} f_{j'}^{(\ell')} = f_i^{(\ell)}. \quad (4.2.37)$$

This means that

Property 4.2.5.2 (Finding SALC). When $P_{ii}^{(\ell)}$ operates on any member of the initial basis, it generates the i -th member of the basis for the irrep of species $\Gamma^{(\ell)}$. Then using $P_{ji}^{(\ell)}$ we can find the j -th member of the set.

However, we normally work with characters, so now we define

$$p^{(\ell)} \equiv \sum_i P_{ii}^{(\ell)} = \frac{d_\ell}{h} \sum_{i,R} D_{ii}^{(\ell)}(R)^* R = \frac{d_\ell}{h} \sum_R \chi^{(\ell)}(R)^* R. \quad (4.2.38)$$

The effect of the operator is

$$p^{(\ell)} f_j = \sum_i P_{ii}^{(\ell)} f_j = \sum_i f_i^{(\ell)}. \quad (4.2.39)$$

This operator thus converts any arbitrary initial basis function into a linear combination of the SALC.

4.2.6 Decomposition of direct product bases

How do we find out the symmetry species spanned by say, xy , if it's not reported in the character table? The more general question is then, given that we know the symmetry species spanned by basis functions (f_1, f_2, \dots) , can we state the symmetry species spanned by their products, such as $(f_1^2, f_1 f_2, \dots)$?

Lemma 4.2.6.1 (Direct-product representation). If $\mathbf{f}^{(\ell)}$ is a basis for an irrep of $\Gamma^{(\ell)}$ of dimension d_ℓ and $\mathbf{f}'^{(\ell')}$ for $\Gamma^{(\ell')}$ of dimension $d_{\ell'}$, then the products (containing mixed terms) also form the basis for a (perhaps reducible) representation, which is called a **direct product representation**, of dimension $d_\ell d_{\ell'}$.

Proof. Under an operation R of a group the basis functions transform as

$$\begin{aligned} R f_i^{(\ell)} &= \sum_j f_j^{(\ell)} D_{ji}^{(\ell)}(R) \\ R f_{i'}^{(\ell')} &= \sum_j f_j^{(\ell')} D_{ji'}^{(\ell')}(R). \end{aligned} \quad (4.2.40)$$

So their product is simply

$$\left(Rf_i^{(\ell)}\right)\left(Rf_{i'}^{(\ell')}\right)=\sum_{j,j'}f_j^{(\ell)}f_{j'}^{(\ell')}D_{ji}^{(\ell)}(R)D_{j'i'}^{(\ell')}(R), \quad (4.2.41)$$

which is a linear combination of the products $f_j^{(\ell)}f_{j'}^{(\ell')}$. □

Now we have the new representation, to reduce it, we just need to work out the character, which can be done really simply:

Theorem 4.2.6.1 (Character of direct-product representation).

$$\chi(r)=\chi^{(\ell)}(R)\chi^{(\ell')}(R) \quad (4.2.42)$$

Proof. The matrix representative of the operation R in the direct-product basis is $D_{ji}^{(\ell)}(R)D_{j'i'}^{(\ell')}(R)$. Note especially that jj' now indexes rows and ii' indexes columns¹. The diagonal elements are the elements with $j=i$ and $j'=i'$. So the character of the operation R in the new basis is

$$\begin{aligned}\chi(r) &= \sum_{i,i'} D_{ii}^{(\ell)}(R)D_{i'i'}^{(\ell')}(R) = \left\{ \sum_i D_{ii}^{(\ell)}(R) \right\} \left\{ \sum_{i'} D_{i'i'}^{(\ell')}(R) \right\} \\ &= \chi^{(\ell)}(R)\chi^{(\ell')}(R)\end{aligned} \quad (4.2.43)$$

□

We can then reduce the characters and get the irreps.

But let's look at this closer: (x, y) can be a basis of the E irrep, and so (x^2, xy, yx, y^2) should span $E \times E = A_1 + A_2 + E$. But since xy and yx are the same function, we should be looking at a 3-degenerate basis, so which one do we discard? We must introduce

¹If the two original bases contain 3 members each, the new basis now must contain 9 members, giving rise to 9x9 representatives. The rows and columns are indexed as 11, 12, 13, 21, 22 and so on.

Property 4.2.6.1 (Symmetrised and antisymmetrised direct product).

$$f_{ij}^{(+)} = \frac{1}{2} \{ f_i^{(\ell)} f_j^{(\ell)} + f_j^{(\ell)} f_i^{(\ell)} \} \quad (4.2.44a)$$

$$f_{ij}^{(-)} = \frac{1}{2} \{ f_i^{(\ell)} f_j^{(\ell)} - f_j^{(\ell)} f_i^{(\ell)} \} \quad (4.2.44b)$$

The characters are given, without proof

$$\chi^{+}(R) = \frac{1}{2} \{ \chi^{(\ell)}(R)^2 + \chi^{(\ell)}(R^2) \} \quad (4.2.45a)$$

$$\chi^{-}(R) = \frac{1}{2} \{ \chi^{(\ell)}(R)^2 - \chi^{(\ell)}(R^2) \} \quad (4.2.45b)$$

So it is clear that the antisymmetrised direct product for xy vanishes, and the basis set $\{0, 0, 0\}$ has a character of $(1, 1, -1)$, which is identified as A_2 .

Therefore the full direct product decomposition is given as

$$E \otimes E = A_1 \oplus [A_2] \oplus E. \quad (4.2.46)$$

So (x^2, xy, y^2) spans $A_1 \oplus E$. Results like this will be important in working out selection rules.

Property 4.2.6.2 (Properties of direct products).

1. The direct product of anything and A , the *totally symmetric irrep*, is itself.
2. The direct produce of anything and itself contains the tot. sym. irrep.

4.2.7 The full rotation group

Full rotation groups are point groups in which rotations through any angles are symmetry operations.

The generators of rotations

Considering the full rotation group R_2 (synonymous to $C_{\infty v}$), the effect on the basis of (x, y) of an infinitesimal counter-clockwise rotation by the angle $\delta\phi$ about the z -axis, as

shown in fig. 4.2 is

$$\begin{aligned}
 C_{\delta\phi}(x, y) &= \{r \cos(\phi - \delta\phi), r \sin(\phi - \delta\phi)\} \\
 &= \{r \cos \phi \cos \delta\phi + r \sin \phi \sin \delta\phi, r \sin \phi \cos \delta\phi - r \cos \phi \sin \delta\phi\} \\
 &= \{r \cos \phi + r \delta\phi \sin \phi + \dots, r \sin \phi - r \delta\phi \cos \phi + \dots\} \\
 &= (x + y \delta\phi + \dots, y - x \delta\phi + \dots) \\
 &= (x, y) - (-y, x) \delta\phi + \dots
 \end{aligned} \tag{4.2.47}$$

For comparison, the effect of the angular momentum operator on (x, y) is

$$\begin{aligned}
 \hat{l}_z(x, y) &= \frac{\hbar}{i} \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) (x, y) \\
 &= \frac{\hbar}{i} (-y, x)
 \end{aligned} \tag{4.2.48}$$

so we can identify

$$C_{\delta\phi} = 1 - \frac{i}{\hbar} \delta\phi \hat{l}_z + \dots \tag{4.2.49}$$

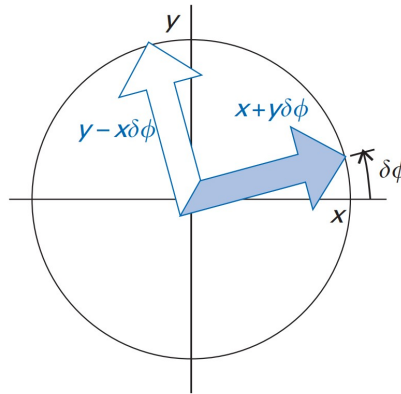


Figure 4.2: The effect of $C_{\delta\phi}$ on the basis of (x, y) .

We have effectively expressed the classical operation $C_{\delta\phi}$ in the language of quantum mechanics, and we can call \hat{l}_z the quantum mechanical generator of rotations about the z -axis²

Now moving onto R_3 , a sequential x and y rotation gives

$$\begin{aligned}
 C_{\delta\beta}^{(y)} C_{\delta\alpha}^{(x)} &= \left(1 - \frac{i}{\hbar} \delta\beta \hat{l}_y + \dots \right) \left(1 - \frac{i}{\hbar} \delta\alpha \hat{l}_x + \dots \right) \\
 &= 1 - \frac{i}{\hbar} (\delta\beta \hat{l}_y + \delta\alpha \hat{l}_x) + \left(\frac{i}{\hbar} \right)^2 \delta\beta \delta\alpha \hat{l}_x \hat{l}_y + \dots
 \end{aligned} \tag{4.2.50}$$

²The classical $\hat{l}_{z,\text{class}}$ does away with the \hbar , i.e., is just $\hat{l}_{z,\text{quant}}/\hbar$

but a sequential y and x rotation gives

$$\begin{aligned} C_{\delta\alpha}^{(x)} C_{\delta\beta}^{(y)} &= \left(1 - \frac{i}{\hbar} \delta\alpha \hat{l}_x + \dots\right) \left(1 - \frac{i}{\hbar} \delta\beta \hat{l}_y + \dots\right) \\ &= 1 - \frac{i}{\hbar} (\delta\beta \hat{l}_y + \delta\alpha \hat{l}_x) + \left(\frac{i}{\hbar}\right)^2 \delta\beta \delta\alpha \hat{l}_y \hat{l}_x + \dots \end{aligned} \quad (4.2.51)$$

so we can see that the difference between these two to the second order is

$$C_{\delta\beta}^{(y)} C_{\delta\alpha}^{(x)} - C_{\delta\alpha}^{(x)} C_{\delta\beta}^{(y)} = \left(\frac{i}{\hbar}\right)^2 \delta\alpha \delta\beta (\hat{l}_y \hat{l}_x - \hat{l}_x \hat{l}_y) = \frac{i}{\hbar} \delta\alpha \delta\beta \hat{l}_z \quad (4.2.52)$$

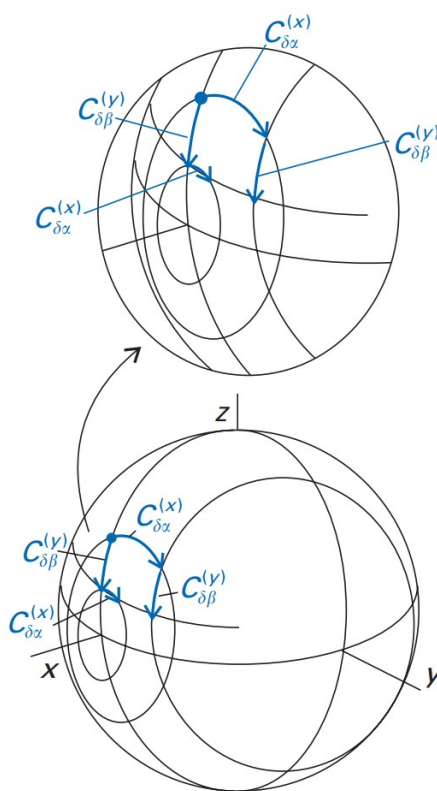


Figure 4.3: The difference between the two composite rotations, we can see the commutation relations graphically.

The representation of the full rotation group

The spherical harmonics Y_{l,m_l} for a given l transform into linear combinations of one another under a rotation, for example, the p -orbitals rotate into one another but not into d -orbitals. This is an exemplification of theorem 4.3.2.1, which says degenerate eigenfunctions can be related by symmetry operations.

Therefore, the functions $Y_{l,l}, Y_{l,l-1}, \dots, Y_{l,-l}$ form a basis for a $(2l+1)$ -dimensional repre-

sensation of the group. As they are orthonormal to each other, the representation is irreducible. The spherical harmonic has the form $Y_{l,m_l} = P_l^{m_l}(\cos \theta) \exp(im_l \phi)$ (eq. 1.4.44) so a rotation by α about the z -axis gives $P(\cos \theta) \exp(im_l(\phi - \alpha))$, extending this to the entire basis we have

$$\begin{aligned}
 & C_\alpha^{(z)}(Y_{l,l}, Y_{l,l-1}, \dots, Y_{l,-l}) \\
 &= P(\cos \theta) \{ \exp[i(l)(\phi - \alpha)], \exp[i(l-1)(\phi - \alpha)], \dots, \exp[-il(\phi - \alpha)] \} \\
 &= (Y_{l,l}, Y_{l,l-1}, \dots, Y_{l,-l}) \begin{bmatrix} e^{-il\alpha} & 0 & 0 & \dots & 0 \\ 0 & e^{-i(l-1)\alpha} & 0 & \dots & 0 \\ 0 & 0 & \ddots & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & e^{il\alpha} \end{bmatrix}
 \end{aligned} \tag{4.2.53}$$

We have identified the matrix representative of the rotation in this basis. The character of rotation is then

$$\begin{aligned}
 \chi(C_\alpha) &= e^{-il\alpha} + e^{-i(l-1)\alpha} + \dots + e^{il\alpha} \\
 &= \frac{e^{-il\alpha}[e^{i(2l+1)\alpha} - 1]}{e^{i\alpha} - 1} \\
 &= \frac{\sin(l + \frac{1}{2})\alpha}{\sin \frac{1}{2}\alpha}
 \end{aligned} \tag{4.2.54}$$

In the limit of $\alpha \rightarrow 0$ (infinitesimal rotation as a symmetry operation) we get the character as $2l + 1$ so we recover the fact that levels with quantum number l are $(2l + 1)$ -fold degenerate in a spherical system.

4.2.8 Group theory and angular momenta

In the last section we saw that a quantum rigid rotor belong to the full rotation group. Now suppose there are two rotors with j_1 and j_2 , belonging the irreps $\Gamma^{(j_1)}$ and $\Gamma^{(j_2)}$ with basis functions $f_{mj_1}^{(j_1)}$ and $f_{mj_2}^{(j_2)}$ respectively.

From section 4.2.6 we know that the bases for $\Gamma^{(j_1)} \otimes \Gamma^{(j_2)}$ are $f_{mj_1}^{(j_1)} f_{mj_2}^{(j_2)}$, subject to further reduction, we can write

$$\Gamma^{(j_1)} \otimes \Gamma^{(j_2)} = \sum_j a_j \Gamma^{(j)} \tag{4.2.55}$$

The character of rotation the system, when viewed through the decoupled picture, is

$$\chi(C_\alpha) = \chi^{(j_1)}(C_\alpha) \chi^{(j_2)}(C_\alpha) = \sum_{m_{j_1}=-j_1}^{j_1} \sum_{m_{j_2}=-j_2}^{j_2} e^{i(m_{j_1}+m_{j_2})\alpha} \tag{4.2.56}$$

As it contains repeated exponentials, we hope to write this in the form of

$$\chi(C_\alpha) = \sum_j \sum_{m_j} e^{im_j \alpha} \quad (4.2.57)$$

which physically represents the coupled picture. To show this is possible, we realise that, purely from the mathematical form, $|m_{j_1} + m_{j_2}| \leq j_1 + j_2$, so $|m_j| \leq j_1 + j_2$ and $j \leq j_1 + j_2$. Therefore $a_j = 0$ for $j > j_1 + j_2$. The maximum m_j can only be created in one way, when $m_{j_1} = j_1$ and $m_{j_2} = j_2$, so $a_{j_1+j_2} = 1$. For $m_j = j_1 + j_2 - 1$, there are two ways to create it but one is already accounted for by the representation with $j = j_1 + j_2$, so $a_{j_1+j_2-1} = 1$, and so on. We now have the series

$$\Gamma^{(j_1)} \otimes \Gamma^{(j_2)} = \Gamma^{(j_1+j_2)} \oplus \Gamma^{(j_1+j_2-1)} \oplus \dots \quad (4.2.58)$$

to terminate, we assume for a moment that $j_1 > j_2$, noting that $\Gamma^{(j_i)}$ is $(2j_i + 1)$ -fold degenerate, so $\Gamma^{(j_1)} \otimes \Gamma^{(j_2)}$ is $(2j_1 + 1)(2j_2 + 1)$ -fold degenerate, if the series terminate at $j_1 - j_2$, the degeneracy on the right hand side equals the left:

$$\begin{aligned} \text{RHS} &= 2[(j_1 + j_2) + (j_1 + j_2 - 1) + \dots + (j_1 + j_2 - 2j_2)] + 2j_2 + 1 \\ &= 4j_1j_2 + 2j_1 + 2j_2 + 1 \\ &= \text{LHS} \end{aligned} \quad (4.2.59)$$

This argument is exactly analogous to that given in section 1.4.5.5. We can write

$$\chi(C_\alpha) = \sum_{j=|j_1-j_2|}^{j_1+j_2} \sum_{m_j=-j}^j e^{im_j \alpha} = \sum_{j=|j_1-j_2|}^{j_1+j_2} \chi^{(j)}(C_\alpha) \quad (4.2.60)$$

This is essentially saying

$$\Gamma^{(j_1)} \otimes \Gamma^{(j_2)} = \Gamma^{(j_1+j_2)} \oplus \Gamma^{(j_1+j_2-1)} \oplus \dots \oplus \Gamma^{(|j_1-j_2|)} \quad (4.2.61)$$

4.3 APPLICATIONS

4.3.1 Vanishing integrals

Consider two functions f and g , one antisymmetric and the other symmetric about $x = 0$. Now consider their integrals on the interval $[-a, +a]$: that of f necessarily zero and that of g can only be zero by accident. Now, if we look closely at the interval $[-a, +a]$ as an object, it belongs to point group C_s : identity and mirror plane. Function f can be a basis of the irrep A'' because $Ef = f$, $\sigma_h f = -f$. Meanwhile, g spans A' , the totally symmetric irrep. So we generalise:

Lemma 4.3.1.1 (Vanishing integral). The integral is necessarily zero if the integrand is a basis for the totally symmetric irreducible representation of the group.

We now examine the products between f and g :

1. f^2 : $A'' \otimes A'' = A'$, non-vanishing.
2. g^2 : $A' \otimes A' = A'$, non-vanishing.
3. fg : $A'' \otimes A' = A''$, vanishing.

Another way to look at this is that f and g are basis functions for different species therefore they must be orthogonal:

Lemma 4.3.1.2 (Generalised vanishing integral). If $f_i^{(\ell)}$ is the i -th member of a basis that spans the irrep of species $\Gamma^{(\ell)}$ of a group, and $f_j^{(\ell')}$ is the j -th member of a basis that spans the irrep of species $\Gamma^{(\ell')}$ of the same group, then for a symmetric range of integration

$$\int f_i^{(\ell)*} f_j^{(\ell')} d\tau \propto \delta_{\ell\ell'} \delta_{ij} \quad (4.3.1)$$

We now are in a position to state the most important result so far

Theorem 4.3.1.1 (Zero integrals). An integral $\int f^{(\ell)*} f^{(\ell')} d\tau$ over a symmetric range is necessarily zero unless the integrand is a basis for the totally symmetric irreducible representation of the group which will be the case only if $\Gamma^{(\ell)} = \Gamma^{(\ell')}$.

This can be extended to integrals of the form

$$I = \int f^{(\ell)*} f^{(\ell')} f^{(\ell'')} d\tau \quad (4.3.2)$$

where the integral is necessarily zero unless the integrand is a basis for the totally symmetric irrep: we take $\Gamma^{(\ell)} \times \Gamma^{(\ell')}$ and expand out, then we take each $\Gamma^{(k)}$ in the expansion and form direct products $\Gamma^{(k)} \times \Gamma^{(\ell')}$. If the tot. sym. irrep does not occur in the final expansion, then the integral is necessarily zero.

An important application of this would be determining whether a matrix element such as $H_{ij} = \langle \psi_i | \psi_j \rangle$ is zero.

In the case of the Hamiltonian, we know it is rotationally invariant *i.e.*, it belongs to the totally symmetric irrep, so all we really need to determine is whether ψ_i and ψ_j belong to the same irrep.

Another, more general example is in the case of the *dipole moment operators*:

In a C_{4v} (SF_5Cl say) molecule, do the integrals $\langle d_{xy} | z | d_{x^2-y^2} \rangle$ and $\langle d_{xy} | l_z | d_{x^2-y^2} \rangle$ vanish? The objects span the following irreps:

| Object | Irrep |
|---------------|-------|
| d_{xy} | B_2 |
| $d_{x^2-y^2}$ | B_1 |
| z | A_1 |
| l_z | A_2 |

Since

$$\begin{aligned} B_2 \otimes A_1 \otimes B_1 &= B_2 \otimes B_1 \neq A_1 \\ B_2 \otimes A_2 \otimes B_1 &= B_2 \otimes B_2 = A_1 \end{aligned} \quad (4.3.3)$$

the former must vanish and the latter not necessarily.

4.3.2 Degeneracy

The Hamiltonian of a system must be invariant under every operation of the point group of the system:

$$RH \equiv H. \quad (4.3.4)$$

A mathematical way of reasoning is that, because $H \equiv T + V$, which, take the one-dimensional harmonic oscillator for example, depends on d^2/dx^2 and x^2 respectively, and

as such it's invariant under replacement of x by $-x$, so the Hamiltonian spans the tot. sym. irrep of C_s .

Because H is invariant under a similarity transformation of the group, *i.e.*, change of basis, or symmetry operation, we can write

$$RHR^{-1} = H. \quad (4.3.5)$$

Right multiplication with R gives $RH = HR$, so we can conclude that symmetry operations must commute with the Hamiltonian. We then introduce

Theorem 4.3.2.1 (Degenerate eigenfunctions). Eigenfunctions that are related by symmetry transformations of the system are degenerate, and *vice versa*.

Proof. Consider $H\psi_i = E\psi_i$. Right multiplying by R gives

$$\begin{aligned} RH\psi_i &= ER\psi_i \\ RH(R^{-1}R)\psi_i &= ER\psi_i \\ HR\psi_i &= ER\psi_i. \end{aligned} \quad (4.3.6)$$

Therefore, ψ_i and $R\psi_i$ are degenerate. □

So how to determine the maximum degree of degeneracy? We recall that the projection operator P_{ij} can give us all members of a basis set, and that it is formed by a linear combination of group operations, so it necessarily commutes with the Hamiltonian, therefore, for an irrep of dimension d ,

$$P_{ij}H\psi_j = HP_{ij}\psi_j = H\psi_i = P_{ij}E\psi_j = E\psi_i. \quad (4.3.7)$$

In the same manner we can generate all d members this way and they are all going to be degenerate, so

Theorem 4.3.2.2 (Degree of degeneracy). The degree of degeneracy of a set of functions is equal to the dimension of the irreducible representation they span, *i.e.*,

$$D = \chi(E). \quad (4.3.8)$$

4.3.3 Chemical bonding

To find out which orbitals participate in bonding in a molecule, first find an appropriate basis and construct a reducible representation of the point group the molecule belongs to, and reduce it. It's that simple. We give a few examples.

σ bonding in D_{3h}

For a choice of basis along the three bonds, we construct the character table

| D_{3h} | E | $2C_3$ | $3C_2$ | σ_h | $2S_3$ | $3\sigma_v$ |
|------------|---|--------|--------|------------|--------|-------------|
| Γ_1 | 3 | 0 | 1 | 3 | 0 | 1 |

Reducing it with the help of D_{3h} character table, we get, for example

$$a(A'_1) = \frac{1}{12}(3 \times 1 + 0 + 1 \times 3 + 3 \times 1 + 0 + 1 \times 3) = 1 \quad (4.3.9)$$

etc., we get

$$\Gamma_1 = A'_1 \oplus E' \quad (4.3.10)$$

A'_1 includes $x^2 + y^2$ and z^2 , which means either the spherically symmetrical s orbital or the d_{z^2} orbital is involved. Same for E' , which either involves p_x and p_y together or $d_{x^2-y^2}$ and d_{xy} together. We know the energy levels of these orbitals and therefore we can say that s , p_x and p_y , i.e., an sp^2 set, is involved in bonding.

π bonding in D_{3h}

The basis vectors relevant for π bonding are a set of three 'out-of-plane' vectors forming the reducible representation Γ_2 and a set of three 'in-plane' vectors forming Γ_3 . The characters and reduction is as follows:

| D_{3h} | E | $2C_3$ | $3C_2$ | σ_h | $2S_3$ | $3\sigma_v$ | irrep |
|------------|---|--------|--------|------------|--------|-------------|--------------------|
| Γ_2 | 3 | 0 | -1 | -3 | 0 | 1 | $A''_2 \oplus E''$ |
| Γ_3 | 3 | 0 | -1 | 3 | 0 | -1 | $A'_2 \oplus E'$ |

Referring to the character table, we can see Γ_2 corresponds to the orbitals of p_z , (d_{xz} , d_{yz}) together and Γ_3 corresponds to either (p_x , p_y) or ($d_{x^2-y^2}$, d_{xy}) together. **what does this really mean? without a priori input how do we know to ignore the d AOs?**

4.3.4 LCAOs

Linear combinations of atomic orbital is a way to form a molecular orbital from a chosen set of basis functions. The symmetry of the LCAO must be the same as the basis functions, and that's where group theory can help us. We use the aromatic cyclopropanium ion as an example.

Cyclopropanium ion - D_{3h}

We choose the p_z atomic orbitals of the carbon atoms as the basis functions and since they have exactly the same directional properties as the out-of-plane vectors, the reduction is just $A_2'' + E''$ again. So we know that the π molecular orbitals include a doubly degenerate pair and one singly degenerate orbital. This is as far as group theory can take us. The relative energies of the orbitals can be supplied by the Huckel theory, which tells us that the energies of the orbitals are $(\alpha + 2\beta)$, $(\alpha - \beta)$, $(\alpha - \beta)$.

Cyclobutadiene - D_{4h}

Again choosing the p_z AOs, we get the reduction of $E_g + A_{2u} + B_{2u}$. The energies from Huckel theory is $(\alpha + 2\beta)$, α , α , $(\alpha - 2\beta)$, and we can identify α with the doubly degenerate E_g .

4.3.5 Molecular orbital correlation diagrams

H₂O

The relevant orbitals are 2H s and O p and s orbitals. The oxygen $p_{x/y/z}$ orbitals transform as A_1 , B_1 , B_2 respectively, the s orbital is just A_1 . The two hydrogen s orbitals, *together*³, transform as $A_1 \oplus B_1$. We can identify the linear combinations that result in the symmetry species easily as $(\phi_1 \pm \phi_2)/\sqrt{2}$ for A_1 and B_1 respectively.

After identifying the symmetry species, we label them on the energy correlation diagram accordingly, but in small caps, as species for specific molecules are conventionally labelled. Only objects that belong to the same symmetry species will interact, producing a higher energy and a lower energy object, belonging to the same species. However the relative energy cannot be predicted by group theory as usual.

³They must transform together since they're in a molecule, similarly, we can only identify *linear combinations* of the two orbitals that transform as a component of the overall symmetry species.

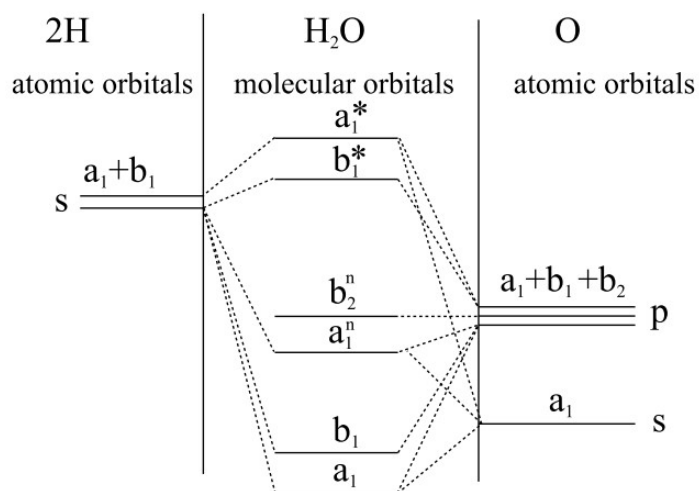


Figure 4.4: The correlation diagram for the water molecule. The two non-bonding orbitals are slightly different in energy, consistent with the photoelectron spectrum.



The relevant *i.e.*, valence orbitals of cobalt are

| | |
|-----------------------------------|----------|
| $3d (x^2 - y^2 \text{ and } z^2)$ | E_g |
| $3d (xy, xz \text{ and } yz)$ | T_{2g} |
| $4s$ | A_{1g} |
| $4p (\text{all})$ | T_{1u} |

And now to reduce the representation formed by the L-M σ bonds, we get $A_{1g} \oplus E_g \oplus T_{1u}$.

Now let the objects belonging to the same species interact, it's clear that the triply degenerate $3d t_{2g}$ orbitals are non-bonding since there are no matching items of the same symmetry.

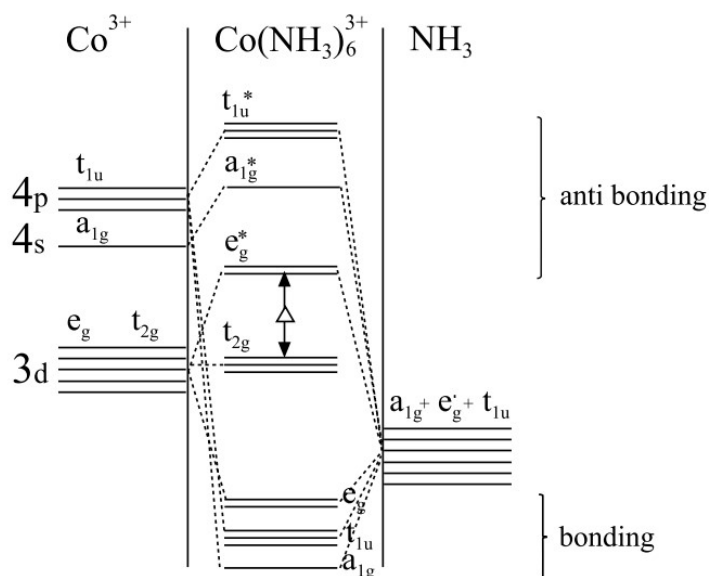


Figure 4.5: The correlation diagram for the cobalt complex. The gap between t_{2g} and e_g^* is the Δ from the Ligand Field Theory.

4.3.6 Molecular vibrations

For a molecule with N atoms there are $3N$ coordinates are needed to specify the spatial arrangements of the atoms. Therefore the molecule has $3N$ free variables that can be independently specified⁴.

It is always true that motions of a body can be decomposed into translations, rotations and internal vibrations. For translations only the coordinates of the CoM needs to be specified, which accounts for 3 degrees of freedom, for rotations an axes need to be specified, this will take up 3 more for a non-linear molecule. For a linear molecule however the axes will not contain a component in the bond axis, therefore only 2 degrees are needed. The remaining are the number of vibrational degrees of freedom.

Theorem 4.3.6.1 (Number of possible vibrations). For a non-linear molecule of N atoms there are $3N - 6$ possible vibrational modes, and $3N - 5$ for a linear molecule.

The questions we will be solving is to use group theory to systematically predict the number of infrared-, and Raman-active stretches for a given molecule. The steps are given

⁴The strict definition of degree of freedom is the number of variables needed to specify the body's configuration in *phase space*, which will include the momentum *i.e.*, velocity vectors, resulting in a total of $6N$ degrees of freedom. Be clear that we're strictly talking about *spatial arrangements* (\mathbb{R}^3) here, without considering the velocity vectors, which in this case does not matter all that much (what's knowing the speed of vibration gonna help with anything? The definition is flexible depending on the problem under consideration).

below:

1. Generate a reducible representation using a basis set of 3 principal axes for every atom
2. Reduce the representation
3. Remove symmetry species corresponding to translations and rotations
4. A vibration (one symmetry species, regardless of its degeneracy) is infrared active if it belongs to the same symmetry species as a component of dipole moment (x , y , or z). It is Raman active if it belongs to the same symmetry species as a component of polarisability (one of the binary products or a combination of them)

H₂O

The 9-vector basis reduces to $3A_1 \oplus A_2 \oplus 3B_1 \oplus 2B_2$, from which we take out the translations A_1 , B_1 and B_2 and rotations A_2 , B_1 and B_2 . The remaining vibrations are $2A_1 \oplus B_1$, which coheres with the 3 degrees of freedom predicted by the rule of thumb.

The vibrations belong to symmetry species of both the dipole and polarisability, therefore water has 3 coincident infrared and Raman bands.

XeF₄

The 15-vector basis reduces to $A_{1g} \oplus A_{2g} \oplus B_{1g} \oplus B_{2g} \oplus E_g \oplus 2A_{2u} \oplus B_{2u} \oplus 3E_u$. After removal of translations and rotations we have $A_{1g} \oplus B_{1g} \oplus B_{2g} \oplus A_{2u} \oplus B_{2u} \oplus 2E_u$, a 9-degenerate representation, which coheres, as it should, with the rule of thumb.

Of the 9 vibrations, A_{1g} , B_{1g} and B_{2g} are Raman active; A_{2u} and $2E_u$ are infrared-active. B_{2u} is active in neither. Therefore there are 3 each, non-coincident infrared and Raman bands for this molecule.

NH₃

The only thing that's different is that the vectors are not pointing in the principal axes directions. This results a little bit of difficulty in the characters of the rotations. The following results are easy to derive:

Property 4.3.6.1 (Character of n -fold rotation and improper rotation).

$$\begin{aligned}\chi(C_n) &= 1 + 2 \cos \frac{2\pi}{n} \\ \chi(S_n) &= -1 + \cos \frac{2\pi}{n}\end{aligned}\tag{4.3.11}$$

H_3^+ The choice of vectors can be made such that three sets of three local axes can be transformed into one another but not mixed *i.e.*, they can be reduced separately. [insert image](#). The vibrations are $A'_1 + E'$, three vibrations as expected from the rule of thumb.

4.3.7 Particular vibrations

If we only need to look at particular vibrations, we can narrow down the choice of basis to just along the vibrations we need, instead of a general $3N$ basis.

For example, if we need to figure out how many carbonyl stretches will be present in a molecule, say $\text{Ni}(\text{CO})_4$ (D_{4h}), only a basis of 4 vectors are needed. It should be noted that the vector here represents extension and compression, not physical displacement from the central metal atom, so they can never be transformed into minus themselves. In this case the basis of 4 vectors transform as $A_{1g} \oplus B_{1g} \oplus E_u$, so we know there will be 1 infrared band and 2 Raman bands. In this way, geometrical isomers can be identified. [fill in more examples when you have the time](#)

Another example is the ethane C-H stretches (or any X-H stretches in other molecules). Under D_{2h} group operations we get $A_g + B_{1g} + B_{2u} + B_{3u}$.

4.3.8 Projection operator method

The projection operator method can be used to determine the linear combination of (an arbitrary set of) basis vectors that makes up a symmetry species. Applications include finding the exact vibrations in a vibrational mode, determining the LCAO, determining the functional form of HAOs, and so on.

The method of the projection operator is outlined as follows [update and link this to the formal approach](#)

1. Choose an arbitrary basis, reduce it
2. Choose a 'generating vector', usually something easy to work with but it's arbitrary. If degeneracy is present in any of the irreps, orthogonal generating vectors has to be chosen (*e.g.* E_g requires two orthogonal generating vectors whereas T_g requires three)
3. Use the *extended* character table, construct the effect of the individual operations on

the generating vector

4. Multiply all the entries with the character of the irrep, sum each row up
5. Normalise the results

Example: Vibrations of BCl_3 (D_{3h})

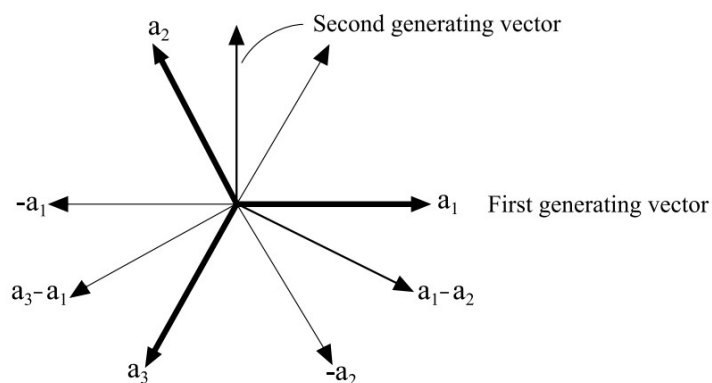


Figure 4.6: The choice of basis and generating vectors.

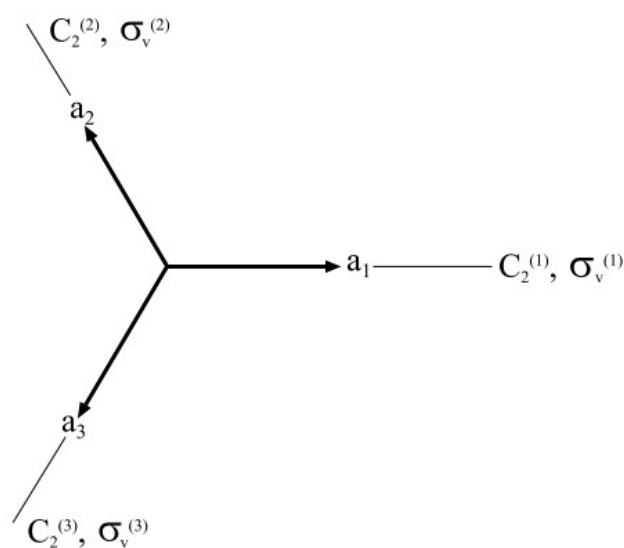


Figure 4.7: The operations.

Step 1: The three-arrow basis along the bonds reduces to $A'_1 + E'$.

Step 2: For the A'_1 mode we use a_1 as the generating vector, and for the E' mode we use a_1 and one vector orthogonal to it, $a_2 - a_3$ ⁵.

Step 3: Using the extended character table, we have

⁵The magnitude doesn't matter.

| D_{3h} | E | C_3 | C_3^2 | $C_2(1)$ | $C_2(2)$ | $C_2(3)$ | σ_h | S_3 | S_3^5 | $\sigma_v(1)$ | $\sigma_v(2)$ | $\sigma_v(3)$ |
|------------------|---------------|-------|---------|----------|----------|----------|------------|-------|---------|---------------|---------------|---------------|
| $A'_1 (a_1)$ | a_1 | a_2 | a_3 | a_1 | a_3 | a_2 | a_1 | a_2 | a_3 | a_1 | a_3 | a_2 |
| $E' (a_1)$ | Same as above | | | | | | | | | | | |
| $E' (a_2 - a_3)$ | ... | | | | | | | | | | | |

Step 4 and 5: Now multiplying with the character and summing the rows, we get the normalised results

| Mode | Orthonormal basis |
|---------|--|
| A'_1 | $\frac{1}{\sqrt{3}}(a_1 + a_2 + a_3)$ |
| $E'(x)$ | $\frac{1}{\sqrt{6}}(2a_1 - a_2 - a_3)$ |
| $E'(y)$ | $\frac{1}{\sqrt{2}}(a_2 - a_3)$ |

It is easy to verify that the basis is orthonormal: this means that all three bonds contribute equally as they must be due to indistinguishability.

The very same approach can be used to find the functional form of the HAOs of boron. We can easily identify the boron s as belonging to A'_1 , p_x and p_y belonging to E' . So we can write

| Orbital | Orthonormal basis |
|---------|--|
| s | $\frac{1}{\sqrt{3}}(a_1 + a_2 + a_3)$ |
| p_x | $\frac{1}{\sqrt{6}}(2a_1 - a_2 - a_3)$ |
| p_y | $\frac{1}{\sqrt{2}}(a_2 - a_3)$ |

Rearranging, we have

| Bond | HAO |
|-------|---|
| a_1 | $\frac{1}{\sqrt{3}}(s + \sqrt{2}p_x)$ |
| a_2 | $\frac{1}{\sqrt{6}}(\sqrt{2}s - p_x + \sqrt{3}p_y)$ |
| a_3 | $\frac{1}{\sqrt{6}}(\sqrt{2}s - p_x - \sqrt{3}p_y)$ |

Again it can be seen that all three orbitals contribute equally to the three HAOs.

Chapter 5

Solid state chemistry

5.1 PRELIMINARY

5.1.1 Bravais lattices

5.1.1.1 Definitions

We establish definitions below.

Definition 5.1.1.1 (Bravais crystal lattice). The *Bravais crystal lattice* is one way to describe crystal lattices, as a translationally periodic array of indistinguishable mathematical points called the *lattice points*. These lattice points are generated by *primitive lattice vectors* (see below).

Definition 5.1.1.2 (Lattice motif or basis). The lattice points are not, in general, the positions of the atoms or molecules that constitute the physical lattice, however, they are imagined to be identically attached to each of the lattice points in a *motif* or *basis*. The motif can be a single atom, a pair of ions (say with one on the lattice point and one a fixed translation away) and so on.

Definition 5.1.1.3 (Lattice vectors). *Lattice vectors* are any vectors connecting one lattice point with another.

Definition 5.1.1.4 (Primitive lattice vectors). *Primitive lattice vectors* can generate the whole lattice by starting at an arbitrary starting lattice point and move by integer multiples of these vectors. These define the edges of *primitive unit cells*.

Definition 5.1.1.5 (Unit cells). Any non-collinear lattice vectors can define the edges of a *unit cell*. These are *not unique*. They are said to tessellate the entire lattice.

Definition 5.1.1.6 (Lattice point count). A unit cell can contain any number of lattice point. They are imagined as 'pies' that are cut by the edges, or wholly contained by the cell. For example, a vertex of a cube on the lattice point will account for $\frac{1}{8}$ of a lattice point and an edge point will account for $\frac{1}{4}$ of a lattice point.

Definition 5.1.1.7 (Primitive unit cells). *Primitive unit cells* are unit cells that contain just one lattice point. These are also *not unique*.

5.1.1.2 Types of Bravais lattices

There are 14 possible Bravais lattices in three dimensions. They are classed first by their lattice systems, which contains requirements for the three lattice angles, the three cell edge lengths (not necessarily primitive lattice vector lengths unless the lattice is primitive), and the symmetry requirements.

Each lattice system can give rise to a maximum of 4 unit cell types:

1. **Primitive (P)**: Unit cell formed by primitive lattice vectors therefore has a lattice point at each corner.
2. **Body-centered (I)**: Primitive with one lattice point at the centre of the cell.
3. **Face-centered (F)**: Primitive with one lattice point at the centre of each face.
4. **Base-centered (C)**: Face-centered but with lattice points at the center of only one pair of opposing faces, which are conventionally oriented to be the base.

We are only going to look at the cubic lattice system, which says $\alpha = \beta = \gamma = 90^\circ$ and $a = b = c$, and requires the symmetry of at least four threefold axes at $109^\circ 28'$ (the tetrahedral angle) to each other. The symmetry requirement excludes the base-centered unit cell type. We have three left to study in detail.

5.1.1.3 The cubic lattices

Simple cubic (cubic P)

The unit cell is just a cube with lattice points at the corners, with side length a , the primitive lattice vectors are

$$\mathbf{a}_1 = a(1, 0, 0) \quad \mathbf{a}_2 = a(0, 1, 0) \quad \mathbf{a}_3 = a(0, 0, 1) \quad V_c = a^3$$

Packing fraction

Imagine spheres centered at the lattice points and they keep expanding until they just touch each other - in this case the spheres on the edges will touch first with $r = \frac{a}{2}$. This means the packing fraction, ρ , of primitive cubic is

$$\rho = \frac{V_{\text{spheres}}}{V_{\text{cell}}} = \frac{\frac{4}{3}\pi r^3}{a^3} = \frac{\frac{4}{3}\pi(\frac{a}{2})^3}{a^3} = \frac{\pi}{6} \approx 52\% \quad (5.1.1)$$

Body-centered cubic (cubic I)

The unit cell that's easiest to picture and work with is a non-primitive unit cell that was described in section 5.1.1.2. The primitive vectors and unit cell volumes are

$$\mathbf{a}_1 = \frac{1}{2}a(1, 1, -1) \quad \mathbf{a}_2 = \frac{1}{2}a(-1, 1, 1) \quad \mathbf{a}_3 = \frac{1}{2}a(1, -1, 1) \quad V_c = \frac{1}{2}a^3$$

Packing fraction The spheres along the body diagonal touch first and is equal to 4 radii. The length of the body diagonal is $\sqrt{3}a = 4r$, so $a = 4r/\sqrt{3}$. There are 2 spheres contained in the unit cell considered, whose volume is a^3 . The packing fraction is then given as

$$\rho = \frac{V_{\text{spheres}}}{V_{\text{cell}}} = \frac{2 \times \frac{4}{3}\pi r^3}{a^3} = \frac{2 \times \frac{4}{3}\pi r^3}{(4r/\sqrt{3})^3} = \frac{\sqrt{3}\pi}{8} \approx 68\% \quad (5.1.2)$$

Face-centered cubic (cubic F)

The unit cell is again non-primitive. The primitive lattice vectors and volume of primitive unit cell are

$$\mathbf{a}_1 = \frac{1}{2}a(1, 1, 0) \quad \mathbf{a}_2 = \frac{1}{2}a(0, 1, 1) \quad \mathbf{a}_3 = \frac{1}{2}a(1, 0, 1) \quad V_c = \frac{1}{4}a^3 \quad (5.1.3)$$

Packing fraction In this case the spheres on the face diagonals touch first and is equal to 4 radii. The length of the face diagonal is $\sqrt{2}a$, so $a = 4r/\sqrt{2}$. The unit cell contains 4

spheres. The packing fraction is therefore

$$\rho = \frac{V_{\text{spheres}}}{V_{\text{cell}}} = \frac{4 \times \frac{4}{3}\pi r^3}{(4r/\sqrt{2})^3} = \frac{\pi}{3\sqrt{2}} \approx 74\% \quad (5.1.4)$$

This is the greatest packing density possible. Therefore the cubic face-centered lattice is also known as *cubic close-packed*.

5.1.1.4 Lattices with a motif

Consider a face-centered cubic lattice, with sodium ions on the lattice points, and chloride ions $(0, 0, \frac{1}{2}a)$ away. A pair of sodium and chloride ions forms the motif for the archetypal NaCl lattice.

5.1.2 The ionic model

5.1.2.1 The model

Consider now a 1-D array of alternating negative and positive charges with z_- and z_+ charges respectively, spaced a apart. Starting arbitrarily from a positive charge and consider it the origin, looking to the right, the net electrostatic interaction is

$$-\frac{z_+z_-e^2}{4\pi\epsilon_0a} + \frac{z_+z_-e^2}{4\pi\epsilon_0(2a)} - \frac{z_+z_-e^2}{4\pi\epsilon_0(3a)} + \dots \quad (5.1.5)$$

Now looking left as well and multiplying by the Avogadro's number we have

$$\begin{aligned} E_{\text{electrostatic}} &= -2N_A \frac{z_+z_-e^2}{4\pi\epsilon_0a} \left(1 - \frac{1}{2} + \frac{1}{3} - \dots\right) \\ &= -2 \ln 2 N_A \frac{z_+z_-e^2}{4\pi\epsilon_0a} \\ &\equiv -\mathcal{A} N_A \frac{z_+z_-e^2}{4\pi\epsilon_0a} \end{aligned} \quad (5.1.6)$$

where $\mathcal{A} = 2 \ln 2$ is the *Madelung constant*, specific to each lattice under consideration.

However this presents the problem that the interaction energy is always attractive and predicts lattice collapse. An *ansatz* repulsion term B/a^n has to be introduced, corresponding to quantum mechanical repulsion between filled orbitals. The order n (typically 9 or greater) is fitted *post-hoc* from experimental data on equilibrium lattice spacing, and the constant B can be eliminated as we will show below. The total energy as a function of

lattice spacing a is written as

$$E(a) = -\frac{N_A \mathcal{A} z_+ z_- e^2}{4\pi\epsilon_0 a} + \frac{B}{a^n} \quad (5.1.7)$$

At equilibrium, $dE/da = 0$ and $a = a_0$, we can write

$$\frac{nB}{a_0^{n+1}} = \frac{N_A \mathcal{A} z_+ z_- e^2}{4\pi\epsilon_0 a_0^2} \quad (5.1.8)$$

which means we can eliminate B by multiplying a_0/n throughout:

$$\frac{B}{a_0^n} = \frac{N_A \mathcal{A} z_+ z_- e^2}{4\pi\epsilon_0 a_0 n} \quad (5.1.9)$$

So the equilibrium energy is

$$E_0 = -\frac{N_A \mathcal{A} z_+ z_- e^2}{4\pi\epsilon_0 a_0} \left(1 - \frac{1}{n}\right) \quad (5.1.10)$$

To get an approximation for lattice energy, which involves the process of



we assume that infinitely separated ions have zero energy, and so $\Delta U = -E_0$, and under such a simplistic model it is acceptable to approximate $\Delta H \approx \Delta U$, so

Theorem 5.1.2.1 (Lattice energy). The lattice energy arising from the ionic model is given as

$$\Delta_L H^\circ \approx \frac{N_A \mathcal{A} z_+ z_- e^2}{4\pi\epsilon_0 a_0} \left(1 - \frac{1}{n}\right) \quad (5.1.11)$$

The Madelung constant for different lattices are given below

| lattice | example | coordination | \mathcal{A} |
|------------------|------------------|--------------|---------------|
| caesium chloride | CsCl | 8:8 | 1.763 |
| fluorite | CaF ₂ | 8:4 | 2.519 |
| rock salt | NaCl | 6:6 | 1.748 |
| rutile | TiO ₂ | 6:3 | 2.408 |
| sphalerite | ZnS | 4:4 | 1.638 |

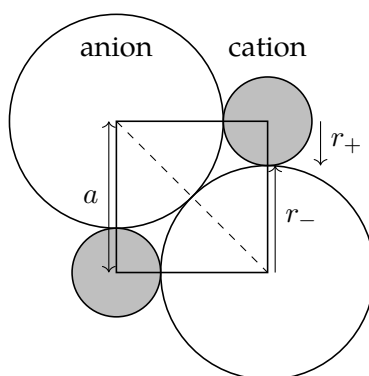
5.1.2.2 Radius ratios

It is not always possible for the ions to touch if say the cations are very large, Cs^+ for example and the anion is very small, F^- for example, the cations will touch first before the anions can touch the cation *optimally*, meaning doing so while it still resides on the motif points. If this happens the lattice energy is not as large as it can be, and a different lattice may be adopted to allow closest approach of the ions.

We can work out the lowest value of r_+/r_- for each lattice type:

Rock salt

We can work out the minimum radius ratio by drawing out the limiting scenario:



in which the diagonal $\sqrt{2}a = 2r_-$, and $a = r_- + r_+$, so we can conclude that

$$\frac{r_+}{r_-} = \frac{a - \sqrt{2}a/2}{\sqrt{2}a/2} = \sqrt{2} - 1 \approx 0.414 \quad (5.1.12)$$

The minimum radius ratios for other lattice types have been worked out:

| lattice | coordination | minimum radius ratio | \mathcal{A} |
|------------------|--------------|----------------------|---------------|
| caesium chloride | 8:8 | 0.732 | 1.763 |
| rock salt | 6:6 | 0.414 | 1.748 |
| sphalerite | 4:4 | 0.225 | 1.638 |

Keeping in mind that $E \propto \mathcal{A}$, the lattice would like to keep the lattice energy as high as possible while maintaining maximum contact. For example, for MgO , $r_+/r_- \approx 0.51$, which is too small for the caesium chloride structure. So it can 'choose' either rock salt

of sphalerite. As rock salt has higher Madelung constant, it should be preferred, and is indeed the case.

5.2 FREE-ELECTRON GAS MODEL

5.2.1 Unconstrained FEG

5.2.1.1 One-dimensional FEG

This model assumes that electrons do not interact with each other or the nuclei. Essentially they are treated as unconstrained gas molecules. In one dimension, the Schrödinger equation was solved in section 1.2.4, with

$$\hat{H} = -\frac{\hbar^2}{2m_e} \frac{d^2}{dx^2} \quad (5.2.1)$$

so the Schrödinger equation reads

$$-\frac{\hbar^2}{2m_e} \frac{d^2\psi}{dx^2} = E\psi \Rightarrow \frac{d^2\psi}{dx^2} = -k^2\psi \quad (5.2.2)$$

where $k = \frac{2m_e E}{\hbar}$ or $E_k = \hbar^2 k^2 / 2m_e$, the time-dependent solution is

$$|\Psi_k(x, t)\rangle = A e^{i(kx - \hbar k^2 t / 2m)} \quad (5.2.3)$$

with momentum

$$\hat{p}|\Psi_k\rangle = \hbar k \quad (5.2.4)$$

The wavelength can be defined as distance before the phase is reset, so in this case

$$\lambda = \frac{2\pi}{|k|} \quad (5.2.5)$$

5.2.1.2 Three-dimensional FEG

In three dimensions, the solution is

$$|\Psi_{\mathbf{k}}(\mathbf{r}, t)\rangle = A e^{i\mathbf{k}\cdot\mathbf{r} - \hbar|\mathbf{k}|^2 t / 2m} \quad (5.2.6)$$

where

$$E_{\mathbf{k}} = \frac{\hbar^2 |\mathbf{k}|^2}{2m_e} \quad (5.2.7)$$

and the momentum is

$$\mathbf{p} = \hbar \mathbf{k} \quad (5.2.8)$$

and the wavelength is

$$\lambda = \frac{2\pi}{|\mathbf{k}|} \quad (5.2.9)$$

Now we need to introduce the important concept of

Definition 5.2.1.1 (*k*-space). As the three-dimensional wavefunctions are defined uniquely by the wavevector \mathbf{k} , these wavevectors can be thought to correspond to a point in the *k*-space. The squared distance from the origin is proportional to energy.

5.2.2 Quantization

The unconstrained systems above are, of course, not quantised, which is a major problem. We must introduce some sort of boundary condition:

Theorem 5.2.2.1 (The Born von Karman boundary conditions). The BvK boundary conditions posits that the two edges of the crystals are continuous, *i.e.*,

$$\Psi(\mathbf{r}, t) = \Psi(\mathbf{r} + L_i \mathbf{a}_i, t) \quad (5.2.10)$$

where L_i is the number of cells in the \mathbf{a}_i direction. In effect, it requires that

$$kL_1 a_1 = 2n_1\pi \Rightarrow k = n_1 \left(\frac{2\pi}{L_1 a_1} \right) \quad (5.2.11)$$

Justification of BvK boundary condition

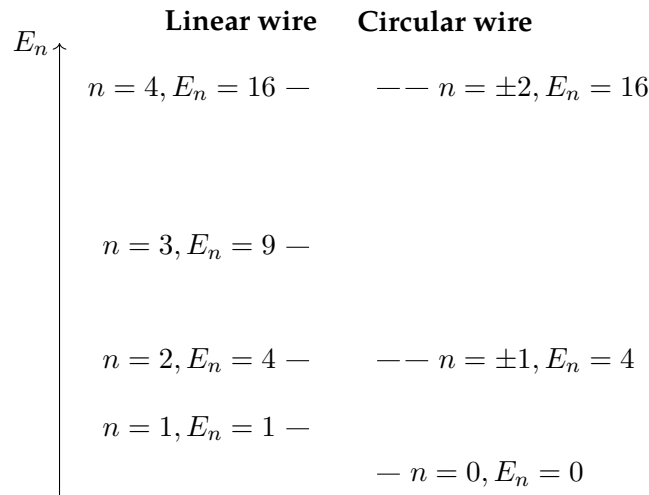
There are two ways to introduce boundary conditions:

- Particle-in-a-box boundary conditions (linear wire)
- BvK boundary conditions (circular wire)

The table below compares the two boundary conditions

| | Linear wire | Circular wire |
|---------------------|-----------------------------------|------------------------------------|
| Boundary conditions | $kL = n\pi, n \in \mathbb{Z}^+$ | $kL = 2n\pi, n \in \mathbb{Z}$ |
| Energies | $E_n = \hbar^2 \pi^2 n^2 / 2mL^2$ | $E_n = 4\hbar^2 \pi^2 n^2 / 2mL^2$ |
| Eigenfunctions | $A \sin(n\pi x / L)$ | $B \exp(2ni\pi x / L)$ |

Basically, the energy levels for the two models look like this



And in the limit that $N \rightarrow \infty$, the density of states is approximately the same, and also the HOMO, now Fermi level, is also the same.

So what's different? The reason for choosing the circular wire is threefold:

1. The eigenfunctions of the linear wire are not eigenfunctions of the momentum operator, while the circular wire ones are, yielding $\hbar \mathbf{k}$ naturally.
2. By ignoring the surface effects, the boundary conditions of the linear wire do not arise as there's no 'outside' where $V(x) = \infty$.
3. The circular wire allows potential energy functions of symmetry $V(x + N\mathbf{a}) = V(x)$, while the linear wire does not. This is important in the full theory.

Simply put, the crystal is imagined to be infinite so that the edge effects are ignored.

In the k -space, the allowed wavefunctions (states) are just a series of points spaced $2\pi/L_1 a_1$ apart. The density of BvK allowed states in 1D is therefore

$$\frac{L_1 a_1}{2\pi} \equiv \frac{\mathcal{L}}{2\pi} \quad (5.2.12)$$

where $\mathcal{L} = L_1 a_1$ is called the BvK length.

In two and three dimensions the density of BvK allowed states are readily generalised to

$$\begin{aligned} \text{2D: } \frac{L_1 L_2 A_c}{(2\pi)^2} &\equiv \frac{\mathcal{A}}{(2\pi)^2} \\ \text{3D: } \frac{L_1 L_2 L_3 V_c}{(2\pi)^3} &\equiv \frac{\mathcal{V}}{(2\pi)^3} \end{aligned} \quad (5.2.13)$$

5.2.3 Fermi-Dirac distribution

The Maxwell-Boltzmann distribution assumes that the particles are distinguishable and any number of particles can occupy a level. However that is clearly not applicable to electrons, which are fermions. The Pauli exclusion principle states that two or more *identical fermions* cannot occupy the same quantum state. A distribution adhering to the Pauli exclusion principle is the

Theorem 5.2.3.1 (Fermi-Dirac distribution). The distribution states that the average number of fermions in a single-particle state i is given by

$$\langle n_i \rangle = \frac{1}{\exp[(\epsilon_i - \mu)/kT] + 1} \quad (5.2.14)$$

where we can write $\langle n_i \rangle$ as $f(\epsilon_i)$, and approximate the chemical potential μ with the *Fermi energy*, ϵ_F , to give

$$f(\epsilon_i) = \frac{1}{\exp[(\epsilon_i - \epsilon_F)/kT] + 1} \quad (5.2.15)$$

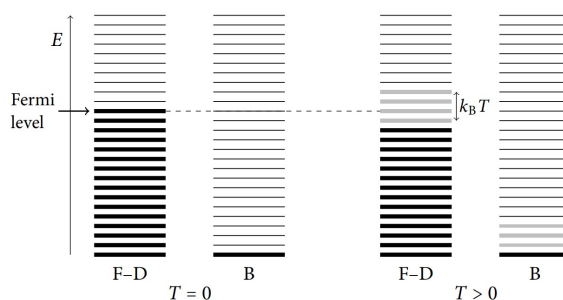


Figure 5.1: Comparison between the predictions of Fermi-Dirac and Maxwell-Boltzmann distributions

The result is a logistic function ranging from 1 for energies much lower than the Fermi level to 0 for energies much higher than the Fermi level. Raising the temperature has the effect of exciting particles near the Fermi level, ‘blurring’ the distribution:

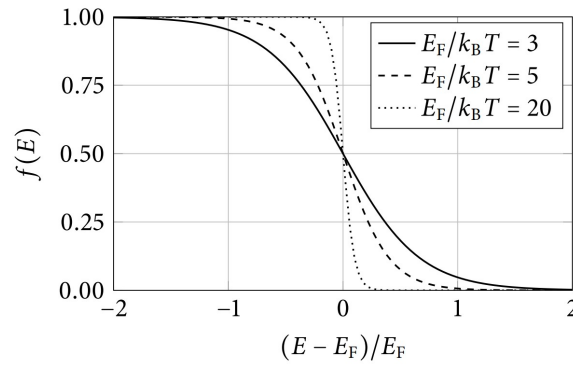


Figure 5.2: The blurring of the Fermi-Dirac distribution occurs at higher temperatures. Notice it always runs from 1 to 0, as required by the Pauli exclusion principle.

Definition 5.2.3.1 (Fermi quantities). The *Fermi energy* is defined only at 0 K, which is defined as the highest filled level at 0K.

The *Fermi level*, E_F is the energy level corresponding to Fermi energy, but remains defined for all temperatures.

The *Fermi wavevector*, k_F is the wavevector corresponding to the Fermi level.

Alternatively, the Fermi level can be approximately seen to be the energy level that has half occupancy at non-zero temperatures.

5.2.4 Properties of the FEG

5.2.4.1 Density of states

The density of BvK states in 3D is

$$W(E) = \underbrace{\frac{4}{3}\pi|\mathbf{k}|^3}_{\text{'volume'}} \times \underbrace{\frac{\mathcal{V}}{(2\pi)^3}}_{\text{density}} \quad (5.2.16)$$

and we know that

$$|\mathbf{k}| = \left(\frac{2m_e E}{\hbar^2} \right)^{\frac{1}{2}} \quad (5.2.17)$$

Keeping in mind that we although each level can accept only one *identical* Fermion, electrons have two spin states and they spin-pair so each level can be doubly occupied, so a

factor of two is applied in front:

$$W(E) = 2 \times \frac{4}{3}\pi|\mathbf{k}|^3 \times \frac{\mathcal{V}}{(2\pi)^3} = \frac{|\mathbf{k}|^3 \mathcal{V}}{3\pi^2} = \frac{\mathcal{V}}{3\pi^2} \left(\frac{2m_e E}{\hbar^2} \right)^{\frac{3}{2}} \quad (5.2.18)$$

Density of states soon follows by simple differentiation

Theorem 5.2.4.1 (Density of BvK states). The number of BvK states with energy up to E is given as

$$W(E) = \frac{\mathcal{V}}{3\pi^2} \left(\frac{2m_e E}{\hbar^2} \right)^{\frac{3}{2}} \quad (5.2.19)$$

so the density of states is given as

$$D(E) = \frac{dW(E)}{dE} = \frac{\mathcal{V}}{2\pi^2} \left(\frac{2m_e}{\hbar^2} \right)^{\frac{3}{2}} E^{\frac{1}{2}} \quad (5.2.20)$$

We note that the density of states increase with energy.

5.2.4.2 Fermi quantities

Suppose that the number density of electrons in the metal is n_e , so there are $n_e \mathcal{V}$ electrons in the BvK volume. At absolute zero, where Fermi energy is defined, we can say that the number of states filled are equal to the number of electrons¹, which is to say

$$W(E_F) = n_e \mathcal{V} \quad (5.2.21)$$

so

$$\begin{aligned} \frac{\mathcal{V}}{3\pi^2} \left(\frac{2m_e E_F}{\hbar^2} \right)^{\frac{3}{2}} &= n_e \mathcal{V} \\ \Rightarrow E_F &= \left(\frac{\hbar^2}{2m_e} \right) (3\pi^2 n_e)^{2/3} \end{aligned} \quad (5.2.22)$$

And the Fermi wavevector can also be worked out:

$$|\mathbf{k}|_F = (3\pi^2 n_e)^{1/3} \quad (5.2.23)$$

To summarise

¹Although we commonly think of the levels as being doubly filled, for the formalism of the BvK we think of two degenerate levels at the same energy, one for the spin up electron one for the spin down.

Theorem 5.2.4.2 (Fermi level and wavevector). For a metal with electron density n_e , the Fermi level and Fermi wavevector are given as

$$E_F = \left(\frac{\hbar^2}{2m_e} \right) (3\pi^2 n_e)^{2/3} \quad (5.2.24)$$

$$|\mathbf{k}|_F = (3\pi^2 n_e)^{1/3}$$

5.2.4.3 Average energy

At absolute zero the average energy is

$$\langle E \rangle = \frac{\text{total energy}}{\text{number of states}} = \frac{\int_0^{E_F} E \times D(E) dE}{\int_0^{E_F} D(E) dE} \quad (5.2.25)$$

The integrals are straightforward and the result is

Theorem 5.2.4.3 (Average energy at absolute zero). The Fermi-Dirac distribution predicts that for a free electron gas, the average energy at absolute zero is

$$\langle E \rangle = \frac{3}{5} E_F \quad (5.2.26)$$

5.2.4.4 Heat capacity

Assuming that the metal atom has only 1 valence electron, one mole of the metal contains N_A valence electrons. To a crude approximation, the number of excited electrons (those within kT of Fermi level) is

$$n = N_A \times \frac{kT}{E_F} = \frac{RT}{E_F} \quad (5.2.27)$$

Again, to an approximation, these electrons on average has kT more in energy², so the total increase in energy as a result of raised temperature is

$$\Delta U = kT \frac{RT}{E_F} \equiv \frac{R}{T_F} T^2 \quad (5.2.28)$$

where $T_F = E_F/k$ is the *Fermi temperature*, so the heat capacity is roughly

$$C_{V,m} = \frac{2R}{T_F} \times T \quad (5.2.29)$$

²They already had some intrinsic kinetic energy as a consequence of the Fermi-Dirac distribution, and now they have an additional kT in 'thermal' energy, which is just additional kinetic energy.

This treatment gives poor agreement with data, as the estimation of the number of electrons is poor and does not take into account the density of states. A more sophisticated treatment gives

$$C_{V,m} = \frac{\pi^2 RT}{2T_F} \quad (5.2.30)$$

5.2.4.5 Bulk modulus

The bulk modulus is the ratio between the pressure change and the fractional change in volume, which, in calculus terms is

$$B = -V \left(\frac{\partial p}{\partial V} \right)_{T,N} \quad (5.2.31)$$

Also the definition of pressure in terms of the internal energy is

$$p = - \left(\frac{\partial U}{\partial V} \right)_N \quad (5.2.32)$$

At absolute zero, we know the average energy, so the internal energy is

$$U = N \langle E \rangle = \frac{3}{5} N E_F \quad (5.2.33)$$

The Fermi energy is

$$E_F = \left(\frac{\hbar^2}{2m_e} \right) (3\pi^2 n_e)^{2/3} \quad (5.2.34)$$

The electron number density, in the formalism of BvK is just

$$n_e = \frac{N}{V} \quad (5.2.35)$$

So we can write

$$U = \frac{3}{5} N \underbrace{\left(\frac{\hbar^2}{2m_e} \right) (3\pi^2 N)^{2/3}}_c \mathcal{V}^{-2/3} \quad (5.2.36)$$

So

$$p = \frac{2}{5} N c \mathcal{V}^{-5/3} \quad (5.2.37)$$

and the bulk modulus is

$$B = \frac{2}{3} N c \mathcal{V}^{-5/3} \quad (5.2.38)$$

Since $E_F = c\mathcal{V}^{-2/3}$, we can finally write

$$B = \frac{2}{3}E_F N/\mathcal{V} = \frac{2}{3}E_F n_e \quad (5.2.39)$$

5.2.4.6 Electrical conductivity

The ‘ground state’ of a 2-dimensional FEG in the k -space at 0 K will show occupied k -states within a circle of radius $|k|$ centred at $(0,0)$, as the highest occupied level is the Fermi level (at 0 K), and the momentum given by $\hbar k$ is on average zero. The circle is called the *Fermi circle*.

Now after the application of an electric field along the x -direction, the average momentum will be in the x -direction³, so the centre of the Fermi circle will be shifted to the right.

It is important to note that the only reason conduction can happen is because there are vacant states just outside the Fermi circle.

A simple model that can give us the conductivity is the

Theorem 5.2.4.4 (The Drude model). The Drude model says that the drift velocity, v_d is given by

$$v_d = e\mathcal{E}\tau/m_e \quad (5.2.40)$$

Proof. The net acceleration experienced by the electron due to the field is $e\mathcal{E}/m_e$. Supposing that the average time between collisions is τ , which knocks the speed down to zero again, the average velocity is just acceleration times time, which gives the result required. \square

The current density is defined as the charge per unit time per unit area, which can be written as

$$j = \frac{n_e v_d \Delta t A e}{\Delta t A} = n_e v_d e = n_e e^2 \mathcal{E} \tau / m_e \quad (5.2.41)$$

The Ohm’s law says that

$$j = \sigma \mathcal{E} \quad (5.2.42)$$

so we can extract

$$\sigma = n_e e^2 \tau / m_e \quad (5.2.43)$$

³It is not infinite because electrons are scattered so a steady-state *drift velocity* is reached.

The *electron mobility*, μ_e (not magnetic susceptibility) is

$$\mu_e \equiv \frac{v_d}{\mathcal{E}} = \frac{e\tau}{m_e} \quad (5.2.44)$$

5.2.5 The nearly-free electron gas

5.2.5.1 Failure of the FEG and Brillouin zones

The free electron gas assumes no interactions whatsoever with the underlying lattice, which is of course not true. The FEG model notably fails when the wavefunctions $\psi_{\mathbf{k}}$ are diffracted by the lattice. To take this into account, the Brillouin zone is defined:

Definition 5.2.5.1 (Brillouin zone). The boundaries of the *Brillouin zone* in k -space are collections of wavefunctions that undergo Bragg diffraction^a. These boundaries are therefore also called Bragg planes. The first Brillouin zone correspond to the collection of k -states reachable from the origin without having to cross a Bragg plane, while the second Brillouin zone are such points reachable by crossing just one Bragg plane, and so on.

^aThis means the waves are refracted by atoms on a series of parallel planes in the *real space*.

We need to find the Bragg planes in k -space now. Referring to **draw diagram**, we can see that the Bragg condition is

$$2a \sin \theta = n\lambda \quad (5.2.45)$$

where n is the integer and will be the order of the Brillouin zone. The de Broglie relation of the incoming waves says

$$\lambda = \frac{h}{|\mathbf{p}|} = \frac{h}{\hbar|\mathbf{k}|} = \frac{2\pi}{|\mathbf{k}|} \quad (5.2.46)$$

So we now have

$$2a \sin \theta = \frac{2\pi n}{|\mathbf{k}|} \quad (5.2.47)$$

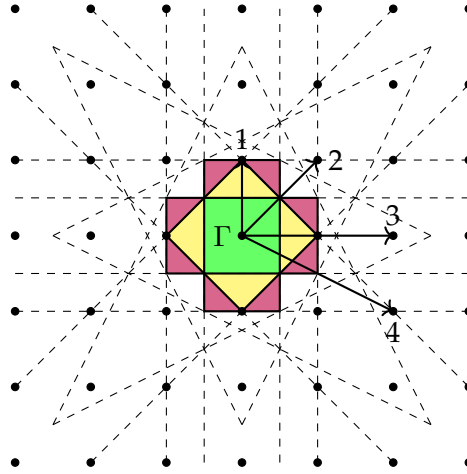
Realising that $|\mathbf{k}| \sin \theta$ is the component of the wavevector perpendicular to the diffracting plane⁴, we can just write

$$k_{\perp} = \frac{n\pi}{a} \quad (5.2.48)$$

In one dimension, as $|\mathbf{k}| = k_{\perp}$, the Bragg planes are just scalar points at integer multiples of $n\pi/a$. In two dimensions, the construction method is to draw perpendicular bisectors connecting the origin and the closest four points in the *reciprocal lattice*, which will enclose

⁴Remember that $\mathbf{p} = \hbar\mathbf{k}$, so the wavevector just lies in the direction of propagation / incidence.

the first Brillouin zone, and then the same is repeated for the next four closest points, for the second Brillouin zone and so on, always extend the Bragg planes as long as possible so you don't get confused whether you've crossed a Bragg plane or not.



Note that the fourth Brillouin zone is not coloured in.

5.2.5.2 Reciprocal lattice

The motivation for defining a reciprocal lattice is to construct a lattice in k -space with a periodicity that's related to the real space lattice. The simplest way to fulfil that is to consider a collection of waves in real space such that

$$\Psi_{\mathbf{k}}(\mathbf{r}) = \Psi_{\mathbf{k}}(\mathbf{r} + \mathbf{R}) \quad (5.2.49)$$

where \mathbf{R} is any lattice vector (not just primitive vectors). Essentially, their phase differ by some integer multiple between 2π between lattice points. A collection of these waves in k -space forms the reciprocal lattice. The equation above can be further developed:

$$\begin{aligned} \Psi_0 e^{i\mathbf{k} \cdot \mathbf{r}} &= \Psi_0 e^{i\mathbf{k} \cdot (\mathbf{r} + \mathbf{R})} \\ \Rightarrow e^{i\mathbf{k} \cdot \mathbf{R}} &= 1 \\ \Rightarrow \mathbf{k} \cdot \mathbf{R} &= 2n\pi \end{aligned} \quad (5.2.50)$$

Considering that $\mathbf{R} = \sum A_i \mathbf{a}_i$ and $\mathbf{k} = \sum B_i \mathbf{b}_i$, we can choose the simplest basis⁵ so that

$$\mathbf{b}_i \cdot \mathbf{a}_j = \delta_{ij} \quad (5.2.52)$$

This way we can construct a reciprocal lattice from a real lattice. The above relation is general and easily applied in 2D cases, however in 3D the following relation, derive from the expression above, can come in handy:

$$\mathbf{b}_3 = \frac{2\pi(\mathbf{a}_1 \times \mathbf{a}_2)}{\mathbf{a}_3 \cdot (\mathbf{a}_1 \times \mathbf{a}_2)} \quad (5.2.53)$$

and cyclic permutations thereof.

The derivation is simple once we note that

$$\frac{(\mathbf{a}_1 \times \mathbf{a}_2)}{\mathbf{a}_3 \cdot (\mathbf{a}_1 \times \mathbf{a}_2)} = \frac{\mathbf{a}_3}{|\mathbf{a}_3|} \quad (5.2.54)$$

We also note that $\mathbf{a}_3 \cdot (\mathbf{a}_1 \times \mathbf{a}_2) = V_c$, the volume of unit cell. The foregoing is summarised below

Theorem 5.2.5.1 (Construction of reciprocal lattice). The reciprocal lattice primitive vectors are given, in general

$$\mathbf{b}_i \cdot \mathbf{a}_j = \delta_{ij} \quad (5.2.55)$$

and in 3D,

$$\mathbf{b}_3 = \frac{2\pi(\mathbf{a}_1 \times \mathbf{a}_2)}{V_c} \quad (5.2.56)$$

and cyclic permutations thereof.

The volume of the reciprocal lattice unit cell is

$$\Omega = \mathbf{b}_1 \cdot (\mathbf{b}_2 \times \mathbf{b}_3) \quad (5.2.57)$$

which is

$$\Omega = \frac{(2\pi)^2}{V_c^3} (\mathbf{a}_1 \times \mathbf{a}_2) \cdot [(\mathbf{a}_2 \times \mathbf{a}_3) \times (\mathbf{a}_3 \times \mathbf{a}_1)] \quad (5.2.58)$$

the big vectorial mess at the end is a number since overall it's a dot product, and since it

⁵The matrix equation is

$$\begin{pmatrix} B_1 & B_2 & B_3 \end{pmatrix} \begin{pmatrix} \mathbf{b}_1 \cdot \mathbf{a}_1 & \mathbf{b}_1 \cdot \mathbf{a}_2 & \mathbf{b}_1 \cdot \mathbf{a}_3 \\ \mathbf{b}_2 \cdot \mathbf{a}_1 & \mathbf{b}_2 \cdot \mathbf{a}_2 & \mathbf{b}_2 \cdot \mathbf{a}_3 \\ \mathbf{b}_3 \cdot \mathbf{a}_1 & \mathbf{b}_3 \cdot \mathbf{a}_2 & \mathbf{b}_3 \cdot \mathbf{a}_3 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} = 2n\pi \quad (5.2.51)$$

so there can be more than one way to choose the basis set $\{\mathbf{b}_i\}$, but obviously the easiest way out is to make it orthogonal to $\{\mathbf{a}_i\}$. The remaining coefficients must have $\sum A_i B_i = n$, since the A_i 's are integers and l is a integer, the B_i 's have to be integers as well, *i.e.*, the \mathbf{b}_i 's are primitive vectors for the reciprocal lattice.

involves every vector twice, it's just V_c^2 , so the reciprocal unit cell volume is

$$\Omega = \frac{(2\pi)^3}{V_c} \quad (5.2.59)$$

5.2.5.3 Boundary cases

The foregoing discussion pointed out the FEG approximations fail at Bragg planes, which is to say the dispersion curve have *discontinuities* at Bragg planes. This provides a motivation for us to study the k -states near the discontinuities more closely.

One way to take the interactions into account is by still looking at the wavefunction predicted by the FEG at the lattice points, and taking that as an indication of how strongly the wavefunction will interact with the lattice at that point.

We note that plane waves with wavenumbers k and $k+n \times (2\pi/a)$ have the same amplitude at all the lattice points spaced a apart. This leads to the idea that the dispersion curve for the entire k -space can be drawn within $-\pi/a$ to $+\pi/a$, as curves stacked on top of each other. This is known as the *reduced zone representation*. **draw diagram**

Now we can add in the interactions. We single out the k -points $\pm\pi/a$, whose corresponding wavefunctions are

$$\psi_+ = \exp(+i\pi x/a) \quad \psi_- = \exp(-i\pi x/a) \quad (5.2.60)$$

These are complex functions, but they are degenerate, so any linear combinations of the two are eigenfunctions too:

$$\psi_c = \frac{1}{2}(\psi_+ + \psi_-) = \cos(\pi x/a) \quad \psi_s = \frac{1}{2i}(\psi_+ - \psi_-) = \sin(\pi x/a) \quad (5.2.61)$$

From **draw diagram**, ψ_s has amplitude 0 at all lattice points and ψ_c has ± 1 (maximum) at all of them. The consequence is that the two eigenfunctions is no longer degenerate, with ψ_c raised in energy. This leads to the creation of a band gap, separating the first and second Brillouin zones.

Predictions about conductivity

The density of BvK states⁶ in one dimension is $La/2\pi$, so in the first Brillouin zone, in the range of $-\pi/a$ to $+\pi/a$, there are L BvK states. So for a metal with one valence eletron,

⁶These can be doubly filled.

the first zone is only half filled. We can visualise it by thinking of a Fermi circle of half the area of the first Brillouin zone, as long as it doesn't touch the Bragg planes, otherwise the circle will be 'flattened' to avoid crossing over the band gap to the second zone.

For metals with two valence electrons, the 1D model will predict that it'll behave like a thermally activated conductor, *i.e.*, a semiconductor, as the band gap must be crossed to produce mobile electrons.

However for a metal with three valence electrons the theory predicts it to be a conductor again, and so on. We obviously don't observe the alternating conductor-semiconductor trend for real metals, and in part it's because we're using the 1D theory for 3D metallic lattices. And from the discussions about Fermi circle above, a Fermi sphere in 3D will behave in even more complicated ways, and quantifying anything becomes increasingly difficult. Clearly, a better, more qualitative theory is needed.

5.3 TIGHT-BINDING MODEL

5.3.1 Bands in one dimension

5.3.1.1 LCAO method

Forming crystal orbitals

Consider a lattice with an orbital suitable for forming bands at each lattice. Using the LCAO method, we create linear combinations, called *crystal orbitals*, written as

$$\psi_k = \sum_{r=1}^N c_k^{(r)} \phi_r \quad (5.3.1)$$

There must be N different crystal orbitals arising from N lattice orbitals, hence the index N . The index r means position.

Boundary conditions

We apply the BvK boundary conditions as before (assuming the wavefunction is continuous between position 1 and position N , *i.e.*, N orbitals form a ring.) The Schrödinger equation is

$$\hat{H}\psi_k = E\psi_k \quad (5.3.2)$$

and the coefficients $c_k^{(r)}$ arise exactly:

$$c_k^{(r)} = e^{ikra}, \quad kNa = 2m\pi, \quad m = 0, \pm 1, \pm 2, \dots \quad (5.3.3)$$

notice that this is essentially the same quantisation we got from the FEG, see Theorem 5.2.2.1. One important thing to note is that, as we switched to the tight-binding model, we can't say $V(x) = 0$, so the Hamiltonian no longer commutes with the momentum operator, so k is no longer proportional to momentum. Instead, k serves as the *symmetry label* for the crystal orbital **supply proof**. But much like the FEG, we can show that all the unique symmetry labels lie within the first Brillouin zone.

Energy calculation with Hückel approximations

The energies of the crystal orbitals are simply

$$E_k = \frac{\langle \psi_k | H | \psi_k \rangle}{\langle \psi_k | \psi_k \rangle} \quad (5.3.4)$$

The Hückel approximations apply to the *atomic orbitals*, whose linear combinations make up ψ_k , with atomic energy α and neighbour interactions β . So the denominator is

$$\begin{aligned} \langle \psi_k | \psi_k \rangle &= \int \left(\sum_{r=1}^N e^{-ikra} \phi_r^* \right) \left(\sum_{r'=1}^N e^{-ikr'a} \phi_{r'} \right) d\tau \\ &= \int \sum_{r=1}^N \phi_r^* \phi_r d\tau = N \end{aligned} \quad (5.3.5)$$

The numerator is

$$\begin{aligned} \langle \psi_k | H | \psi_k \rangle &= \int \left(\sum_{r=1}^N e^{-ikra} \phi_r^* \right) H \left(\sum_{r'=1}^N e^{-ikr'a} \phi_{r'} \right) d\tau \\ &= \sum_{r,r'} \int e^{ik(r'-r)a} \phi_r^* H \phi_{r'} d\tau \end{aligned} \quad (5.3.6)$$

Part II

Thermodynamics and statistical mechanics

Chapter 6

Thermodynamics

6.1 ENTROPY AND THE BOLTZMANN LAW

This section is adapted from Chapters 2 and 5 of [8].

Definition 6.1.0.1 (Equilibrium). An equilibrium defines where a system tends to go and stay, *i.e.*, the net force *analogue* on the system is zero.

Definition 6.1.0.2 (Extremum principles). An *extremum principle* is a principle stating that a macrostate with most number of microstates are the most probable macrostate.

Postulate 6.1.0.1 (Statistical definition of entropy). The statistical definition of entropy, which follows from a *principle of fair apportionment*, states that

$$S = k \ln(W) \quad (6.1.1)$$

We will be showing the intuitions behind Postulate 6.1.0.1 shortly, but we will first show that the alternative definition that

$$\frac{S}{k} = - \sum_{i=1}^t p_i \ln(p_i) \quad (6.1.2)$$

is equivalent to Equation 6.1.1: Roll a t -sided die N times. The multiplicity of outcomes is given by

$$W = \frac{N!}{n_1! n_2! \cdots n_t!}, \quad (6.1.3)$$

where n_i is the number of times side i appears face up.

Stirling's approximation gives that

$$x! \approx \left(\frac{x}{e}\right)^x, \quad (6.1.4)$$

and expressing probabilities $p_i = n_i/N$, we have

$$\begin{aligned} W &= \frac{(N/e)^N}{(n_1/e)^{n_1} (n_2/e)^{n_2} \cdots (n_t/e)^{n_t}} \\ &= \frac{1}{p_1^{p_1} p_2^{p_2} \cdots p_t^{p_t}} \\ \therefore \ln(W) &= - \sum_{i=1}^t n_i \ln p_i \\ \Rightarrow \frac{1}{N} \ln W &= - \sum_{i=1}^t p_i \ln p_i = \frac{S_N}{Nk} = \frac{S}{k}, \end{aligned} \quad (6.1.5)$$

where S_N is the total entropy for N trials.

The **central problem** of statistical thermodynamics is to infer the probability distribution from an observable quantity.

Theorem 6.1.0.1 (Flat distribution from no constraints). The principle of maximising entropy leads to a flat probability distribution when there are no constraints.

Proof. This is essentially a lagrange multiplier problem with

$$\text{objective function: } S(p_1, p_2, \dots, p_t) = -k \sum_{i=1}^t p_i \ln p_i \quad (6.1.6a)$$

$$\text{constraint function: } \sum_{i=1}^t p_i = 1 \Rightarrow \sum_{i=1}^t dp_i = 0 \quad (6.1.6b)$$

We hence construct the multiplier function

$$L(p_i, \lambda) \equiv S - \lambda \left[\left(\sum_{i=1}^t p_i \right) - 1 \right] \quad (6.1.7)$$

Maximising requires

$$\nabla L = 0 \quad (6.1.8)$$

This implies

$$-1 - \ln(p_i) - \lambda = 0 \Rightarrow p_i^* = e^{-1-\lambda}, \quad (6.1.9)$$

where $*$ indicates maximised. We observe that this is a flat probability distribution (sub-

ject to normalisation).

□

Theorem 6.1.0.2 (Boltzmann distribution). However, if there is bias/constraint present, we can show that an exponential distribution results.

Proof. The problem is now summarised as follows

1. A generalised die has t sides,
2. A 'score' ϵ_i is associated with the i -th side,
3. The only observable quantity is

$$\langle \epsilon \rangle \equiv \sum_{i=1}^t p_i \epsilon_i. \quad (6.1.10)$$

We seek to find the distribution of p_i^* . Again this is a lagrange multiplier problem with

$$\text{objective function: } S(p_1, p_2, \dots, p_t) = -k \sum_{i=1}^t p_i \ln p_i \quad (6.1.11a)$$

$$\text{constraint function 1: } g(p_i) \equiv \sum_{i=1}^t p_i = 1 \quad (6.1.11b)$$

$$\text{constraint function 2: } h(p_i) \equiv \sum_{i=1}^t p_i \epsilon_i = \langle \epsilon \rangle \quad (6.1.11c)$$

We again require that

$$\nabla L \equiv \nabla (S - \lambda g - \mu h) = 0 \quad (6.1.12)$$

We have

$$-1 - \ln(p_i^*) - \lambda - \mu \epsilon_i = 0 \quad (6.1.13a)$$

$$p_i^* = e^{-1-\lambda-\mu \epsilon_i} \quad (6.1.13b)$$

$$p_i^* = \frac{p_i^*}{\sum_{i=1}^t p_i^*} = \frac{e^{-\mu \epsilon_i}}{\sum_{i=1}^t e^{-\mu \epsilon_i}}, \quad (6.1.13c)$$

We can express $\langle \epsilon \rangle$ in terms of the distribution from Equation 6.1.11c:

$$\langle \epsilon \rangle = \sum_{i=1}^t \epsilon_i p_i^* = \frac{1}{q} \sum_{i=1}^t \epsilon_i e^{-\mu \epsilon_i}. \quad (6.1.14)$$

□

Definition 6.1.0.3 (Boltzmann distribution law and partition function). Equation 6.1.13c is known as the **Boltzmann distribution law**, and the quantity in the denominator is known as the **partition function** q .

Worked example 6.1.0.1 (Bias in dice). Suppose we have a 6-faced die and the score on each face is the index, *i.e.*, $\epsilon_i = i$.

The Boltzmann distribution law of Equation 6.1.13c gives

$$p_i^* = \frac{e^{-\mu i}}{\sum_{i=1}^6 e^{-\mu i}}. \quad (6.1.15)$$

And from Equation 6.1.14 we have

$$\begin{aligned} \langle \epsilon \rangle &= \sum_{i=1}^6 i p_i^* \\ &= \frac{e^{-\mu} + 2e^{-2\mu} + 3e^{-3\mu} + 4e^{-4\mu} + 5e^{-5\mu} + 6e^{-6\mu}}{e^{-\mu} + e^{-2\mu} + e^{-3\mu} + e^{-4\mu} + e^{-5\mu} + e^{-6\mu}}. \end{aligned} \quad (6.1.16)$$

Equation 6.1.16 is an equation with one unknown, μ . Simple algebra can show that if $\langle \epsilon \rangle = 3.5$, $\mu = 0$ and a flat distribution is resulted, and if $\langle \epsilon \rangle < 3.5$, $\mu < 0$ and *vice versa*.

6.2 THE FUNDAMENTAL THERMODYNAMIC EQUATIONS

This section is adapted from [8] unless otherwise stated.

Theorem 6.2.0.1 (Fundamental equations). The **fundamental thermodynamic equations for entropy and energy** are

$$S = S(U, V, \mathbf{N}) \quad (6.2.1a)$$

$$U = U(S, V, \mathbf{N}) \quad (6.2.1b)$$

respectively, where \mathbf{N} means N_1, N_2, \dots, N_M . These equations *predict equilibria*.

Both equations can completely specify the state of a simple system. These equations does *not* tell you the mathematical dependences, instead, *equations of state*, such as the ideal gas law, specify the interrelations among these variables. These must come from experiments or microscopic models.

6.2.1 The fundamental equations

Taking the total differential of S and U we have

$$dS = \left(\frac{\partial S}{\partial U} \right)_{V, \mathbf{N}} dU + \left(\frac{\partial S}{\partial V} \right)_{U, \mathbf{N}} dV + \sum_{j=1}^M \left(\frac{\partial S}{\partial N_j} \right)_{U, V, N_{i \neq j}} dN_j \quad (6.2.2a)$$

$$dU = \left(\frac{\partial U}{\partial S} \right)_{V, \mathbf{N}} dS + \left(\frac{\partial U}{\partial V} \right)_{S, \mathbf{N}} dV + \sum_{j=1}^M \left(\frac{\partial U}{\partial N_j} \right)_{S, V, N_{i \neq j}} dN_j \quad (6.2.2b)$$

From these expressions we make the following definitions:

Definition 6.2.1.1 (Basic thermodynamic quantities).

$$\text{the temperature: } T = \left(\frac{\partial U}{\partial S} \right)_{V, N} \quad (6.2.3a)$$

$$\text{the pressure: } p = - \left(\frac{\partial U}{\partial V} \right)_{S, N} \quad (6.2.3b)$$

$$\text{the chemical potential: } \mu_j = \left(\frac{\partial U}{\partial N_j} \right)_{S, V, N_{i \neq j}} \quad (6.2.3c)$$

We now have, by way of definition,

$$dU = T dS - p dV + \sum_{j=1}^M \mu_j dN_j \quad (6.2.4a)$$

$$dS = \left(\frac{1}{T} \right) dU + \left(\frac{p}{T} \right) dV - \sum_{j=1}^M \left(\frac{\mu_j}{T} \right) dN_j, \quad (6.2.4b)$$

where Equation 6.2.4b is obtained by algebraic rearrangement of Equation 6.2.4a and gives us three further definitions:

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U} \right)_{V, N} \quad (6.2.5a)$$

$$\frac{p}{T} = \left(\frac{\partial S}{\partial V} \right)_{U, N} \quad (6.2.5b)$$

$$\frac{\mu_j}{T} = \left(\frac{\partial S}{\partial N_j} \right)_{U, V, N_{i \neq j}}. \quad (6.2.5c)$$

We will show that Equations 6.2.4a and 6.2.4b can be used to identify states of equilibrium shortly. In the following section we first show the derivation of the ideal gas law.

6.2.2 The ideal gas law

Theorem 6.2.2.1 (The ideal gas law). The ideal gas law is stated as

$$pV = NkT. \quad (6.2.6)$$

Proof. Imagine a lattice of M sites for N particles where $M \geq N$, the multiplicity W is computed by

$$W = \frac{M!}{N!(M-N)!}. \quad (6.2.7)$$

Hence the entropy is

$$S = k \ln(W(N, M)) = \ln \left[\frac{M!}{N!(M-N)!} \right]. \quad (6.2.8)$$

We now in effect know how the entropy depends on volume, and that's handy because Equation 6.2.5b gives us a way to get an expression for p .

In our model, we assert that $V/M = v_0$, the volume per lattice site, which is a constant. We can now write

$$\left(\frac{\partial S}{\partial V} \right)_N = \left(\frac{dS}{dM} \right)_N \left(\frac{dM}{dV} \right) = \left(\frac{\partial S}{\partial M} \right)_N \left(\frac{1}{v_0} \right). \quad (6.2.9)$$

We still have $(\partial S / \partial M)_N$ to calculate:

$$\frac{S}{k} = \ln(W(N, M)) = \ln \left[\frac{M!}{N!(M-N)!} \right] \quad (6.2.10a)$$

$$= M \ln M - N \ln N - (M - N) \ln(M - N) \quad (6.2.10b)$$

$$= -N \ln \left(\frac{N}{M} \right) - (M - N) \ln \left(\frac{M - N}{M} \right) \quad (6.2.10c)$$

$$\Rightarrow \left(\frac{\partial S}{\partial M} \right)_N = k \left[1 + \ln M - \ln(M - N) - \frac{M - N}{M - N} \right] \quad (6.2.10d)$$

$$= -k \ln \left(1 - \frac{N}{M} \right), \quad (6.2.10e)$$

where Equation 6.2.10d results from rewriting $(M \ln M)$ into $(N \ln M) + [(M - N) \ln M]$. We perform a Taylor series expansion on Equation 6.2.10e, in the limit that $N/M \ll 1$:

$$p = \frac{T}{v_0} \left(\frac{\partial S}{\partial M} \right)_N \quad (6.2.11a)$$

$$= -kT \left(\frac{M}{V} \right) \ln \left(1 - \frac{N}{M} \right) \quad (6.2.11b)$$

$$\approx \left(-\frac{MkT}{V} \right) \left(-\frac{N}{M} \right) \left[1 + \frac{1}{2} \left(\frac{N}{M} \right) + \dots \right] \quad (6.2.11c)$$

$$\approx \frac{NkT}{V}, \quad (6.2.11d)$$

which yields the ideal gas law. Higher-order terms give further refinements, such as the *van der Waals equation*, which includes the first-order correction. \square

Corollary 6.2.2.1 (Entropy of ideal gas). As given by Equations 6.2.6 and 6.2.5b, we can write, for an ideal gas,

$$\left(\frac{\partial S}{\partial V}\right)_{U,N} = \frac{Nk}{V} \quad (6.2.12a)$$

$$S(V, N) = Nk \ln V. \quad (6.2.12b)$$

6.2.3 Second law and equilibria

Definition 6.2.3.1 (Extensive and intensive quantities). *Extensive quantities* are additive and depend on the size of the system, such as internal energy, entropy, volume and number of particles.

Intensive quantities are not additive and do not depend on the size of the system, such as temperature and pressure.

Definition 6.2.3.2 (First-order homogeneous function). Since $U = U(S, V, N)$ is a function of all extensive quantities, U itself is an extensive quantity as well. Furthermore, U has the important property of being a *first-order homogeneous function*, which behaves as

$$f(\lambda x) = \lambda f(x). \quad (6.2.13)$$

This leads to the important property:

Property 6.2.3.1 (Derivative of first-order homogeneous functions). Because, for a first-order homogeneous function,

$$\frac{df(\lambda x)}{d(\lambda x)} = \frac{d[\lambda f(x)]}{dx} \frac{dx}{d(\lambda x)} = \frac{df}{dx}, \quad (6.2.14)$$

i.e., its derivative with respect to an extensive variable is an *intensive variable*.

Therefore, we can conclude from Equations 6.2.3a to 6.2.3c that T , p and μ are intensive quantities and do not depend on system size.

6.2.3.1 The second law

Postulate 6.2.3.1 (The second law). The second law states that isolated systems tend toward their states of maximum entropy.

We will show that this predicts states of equilibrium.

Theorem 6.2.3.1 (Thermodynamic forces). It can be shown that $1/T$ is a measure of a system's tendency for heat exchange, p/T is a measure of a system's tendency for volume exchange, and μ/T , particle exchange.

The strategy to do so is to

1. Determine the relevant independent variables,
2. Apply the fundamental entropy equation and the constraints,
3. Use the maximum entropy principle, *i.e.*, the second law to specify the state of equilibrium.

Temperature. We need to demonstrate that temperature as defined by Equation 6.2.3a satisfies the notion of temperature given as:

- It is the quantity that tells you when heat exchange will occur, and
- objects exchange heat to reach a state of maximum entropy, and
- in this state their temperatures are equal.

Imagine two objects A and B are brought into thermal contact. They can only exchange energy with each other not the surroundings, and no particle or volume exchange is possible either. We are interested to know how the entropy of the objects depend on the energy: $S_A(U_A)$ and $S_B(U_B)$.

Say object A has energy U_A , entropy S_A and $1/T_A = (\partial S_A / \partial U_A)$ and likewise for B .

Entropy is an extensive property, so

$$S_{total} = S_A + S_B. \quad (6.2.15)$$

Equation 6.2.15 does not mean that S_{total} is fixed however. We obtain an equation likewise for total internal energy

$$U_{total} = U_A + U_B = \text{constant}, \quad (6.2.16)$$

which is the constraint equation. To find the state of thermal equilibrium, we apply the second law, *i.e.*,

$$dS = 0. \quad (6.2.17)$$

We therefore have, for $dV = dN = 0$,

$$dS_{total} = dS_A + dS_B = \left(\frac{\partial S_A}{\partial U_A} \right)_{V,N} dU_B + \left(\frac{\partial S_B}{\partial U_B} \right)_{V,N} dU_B = 0. \quad (6.2.18)$$

The differential form of the constraint equation Equation 6.2.16 is

$$\begin{aligned} dU_A + dU_B &= dU_{total} = 0 \\ \Rightarrow dU_A &= -dU_B \end{aligned} \quad (6.2.19)$$

Substitute Equation 6.2.19 into Equation 6.2.18 and rearrange to get

$$dS_{total} = \left[\left(\frac{\partial S_A}{\partial U_A} \right)_{V,N} - \left(\frac{\partial S_B}{\partial U_B} \right)_{V,N} \right] dU_A = 0 \quad (6.2.20a)$$

$$\Rightarrow \left(\frac{\partial S_A}{\partial U_A} \right)_{V,N} = \left(\frac{\partial S_B}{\partial U_B} \right)_{V,N} \quad (6.2.20b)$$

$$\Rightarrow \frac{1}{T_A} = \frac{1}{T_B} \Rightarrow T_A = T_B. \quad (6.2.20c)$$

□

Pressure. Consider a cylinder partitioned into subsystems A and B , with only volume exchange possible. This also means that $T_A = T_B$ since no energy exchange is possible. The constant volume and internal energy constraints leads to

$$dV_A = dV_B \quad (6.2.21a)$$

$$dU_A = dU_B. \quad (6.2.21b)$$

The second law gives

$$dS = \left(\frac{\partial S_A}{\partial V_A} \right) dV_A + \left(\frac{\partial S_B}{\partial V_B} \right) dV_B + \left(\frac{\partial S_A}{\partial U_A} \right) dU_A + \left(\frac{\partial S_B}{\partial U_B} \right) dU_B = 0 \quad (6.2.22)$$

Applying constraints in Equation 6.2.21 gives

$$dS = \left(\frac{p_A}{T_A} - \frac{p_B}{T_B} \right) dV_A + \left(\frac{1}{T_A} - \frac{1}{T_B} \right) dU_A = 0, \quad (6.2.23)$$

Because $T_A = T_B$, we get, at equilibrium,

$$p_A = p_B. \quad (6.2.24)$$

□

A statistical mechanical proof of pressure. Consider a container of M sites with M_A and M_B sites in the two sides. N_A and N_B are fixed. We want to find M_A^* and M_B^* that maximises the multiplicity function:

$$W(N, M) = \frac{M!}{(M - N)!} \frac{1}{N!}. \quad (6.2.25)$$

For $N/M \ll 1$, the approximation

$$\frac{M!}{(M - N)!} \approx M^N \quad (6.2.26)$$

holds (nb. this is a rather trivial result of the limit and *not* Stirling's approximation). Setting $M_B = M - M_A$, Equation 6.2.25 can be reduced to

$$W_{tot} = W_A W_B = \frac{M_A^{N_A} (M - M_A)^{N_B}}{N_A! N_B!} \quad (6.2.27)$$

We are interested to know maximise W in the form of the function $S = k \ln(W)$ with respect to the M 's:

$$\frac{d \ln W}{dM_A} = 0. \quad (6.2.28)$$

Equation 6.2.27 gives us

$$\ln W = N_A \ln M_A + N_B \ln(M - M_A), \quad (6.2.29)$$

and taking logarithm we get

$$\frac{N_B}{M_B^*} = \frac{N_A}{M_A^*}, \quad (6.2.30)$$

which is the desired result that equilibrium is reached when densities are equalised at low pressures. \square

Chemical potential. This proof is adapted from [7] as the argument given in [8] contains omissions.

Consider N identical particles, divided into N_A and N_B . Particle and energy exchange is allowed, but not volume. We thus have constraints

$$dN_A + dN_B = 0 \quad (6.2.31a)$$

$$dU_A + dU_B = 0. \quad (6.2.31b)$$

The second law requires

$$dS = \left(\frac{1}{T_A} - \frac{1}{T_B} dU_A \right) - \left(\frac{\mu_A}{T_A} - \frac{\mu_B}{T_B} \right) dN_A = 0 \quad (6.2.32)$$

We have the conditions of equilibrium

$$\frac{1}{T_A} = \frac{1}{T_B} \quad (6.2.33a)$$

$$\frac{\mu_A}{T_A} = \frac{\mu_B}{T_B}. \quad (6.2.33b)$$

The conditions can be simplified if $T_A = T_B = T$, to

$$dS = \frac{\mu_B - \mu_A}{T} dN_A. \quad (6.2.34)$$

And since $dS \geq 0$, we conclude that if $\mu_B < \mu_A$, dN_A must be negative and hence matter flows from A to B . □

6.3 THE FIRST AND SECOND LAWS IN DEPTH

This section is adapted from Chapters 11 through 14 of [5].

6.3.1 The First Law

6.3.1.1 Energy

The first law is essentially a statement of conservation of energy.

Definition 6.3.1.1 (First Law). Energy is conserved and heat and work are both forms of energy.

Conservation of energy reasons that

$$dU = dQ + dW. \quad (6.3.1)$$

Only in a reversible process, *i.e.*, no turbulence, can we assert that

$$dW = -p dV \quad (6.3.2)$$

Definition 6.3.1.2 (Reversible Process). A reversible process is a process whose direction can be changed by an infinitesimal change in a system parameter. These processes are called **quasistatic**.

It turns out that physically internal energy U is a function of temperature T and volume V , *i.e.*, $U \equiv U(T, V)$. Therefore,

$$dU = \left(\frac{\partial U}{\partial T} \right)_V dT + \left(\frac{\partial U}{\partial V} \right)_T dV \quad (6.3.3)$$

Rearranging Equation 6.3.1 with Equation 6.3.2 we obtain, for *reversible* processes

$$\begin{aligned} dQ &= \left(\frac{\partial U}{\partial T} \right)_V dT + \left[\left(\frac{\partial U}{\partial V} \right)_T + p \right] dV \\ \frac{dQ}{dT} &= \left(\frac{\partial U}{\partial T} \right)_V + \left[\left(\frac{\partial U}{\partial V} \right)_T + p \right] \frac{dV}{dT} \end{aligned} \quad (6.3.4)$$

Definition 6.3.1.3 (Heat capacities).

$$C_V = \left(\frac{\partial Q}{\partial T} \right)_V \quad (6.3.5a)$$

$$C_p = \left(\frac{\partial Q}{\partial T} \right)_p \quad (6.3.5b)$$

From Equation 6.3.4 and Definition 6.3.1.3, we identify

Theorem 6.3.1.1 (Heat Capacities).

$$C_V = \left(\frac{\partial U}{\partial T} \right)_V \quad (6.3.6a)$$

$$C_p = \left(\frac{\partial U}{\partial T} \right)_V + \left[\left(\frac{\partial U}{\partial V} \right)_T + p \right] \left(\frac{\partial V}{\partial T} \right)_p. \quad (6.3.6b)$$

Corollary 6.3.1.1 (Monatomic Heat Capacities). For an ideal monatomic gas, U is given by $U = \frac{3}{2}RT$ per mole. Hence,

$$\left(\frac{\partial U}{\partial V} \right)_T = 0 \quad (6.3.7)$$

and according to the ideal gas law $V = \frac{RT}{p}$,

$$\left(\frac{\partial V}{\partial T} \right)_p = \frac{R}{p}. \quad (6.3.8)$$

Therefore we have, from Equations 6.3.6a and 6.3.6b,

$$C_{V,m} = \frac{3}{2}R \quad (6.3.9a)$$

$$C_{p,m} = \frac{5}{2}R. \quad (6.3.9b)$$

Definition 6.3.1.4 (Adiabatic Index).

$$\gamma \equiv \frac{C_p}{C_V} \quad (6.3.10)$$

For a monatomic gas,

$$\gamma = \frac{C_p}{C_V} = \frac{C_V + R}{C_V} = 1 + \frac{R}{C_V} = \frac{5}{3}. \quad (6.3.11)$$

Corollary 6.3.1.2 (U per unit mass and volume for Ideal Gases). The internal energy of an ideal gas is given by $U = C_v T$, using $pV = Nk_B T$ and $\rho = Nm/V$, we get

$$\frac{p}{\rho} = \frac{k_B T}{m}, \quad (6.3.12)$$

and from Equation 6.3.11, $C_V = \frac{R}{\gamma-1}$, gives

$$U = C_V T = \frac{RT}{\gamma-1} = \frac{N_A k_B T}{\gamma-1}. \quad (6.3.13)$$

As molar mass is mN_A , dividing through by molar mass gives

$$\tilde{u} = \frac{p}{\rho(\gamma-1)}, \quad (6.3.14)$$

the internal energy per unit mass. It follows that the internal energy per unit volume is then simply

$$u = \rho \tilde{u} = \frac{p}{\gamma-1} \quad (6.3.15)$$

As this is a reversible change, since we have $\Delta S = \Delta Q/T$, we can write the molar entropy change

$$\Delta S = R \ln \frac{V_f}{V_i}. \quad (6.3.16)$$

6.3.1.2 Isothermal and Adiabatic Processes

Theorem 6.3.1.2 (Isothermal expansion heat). In an isothermal process $dT \equiv 0$, hence $dU = 0$. This implies

$$\delta W = -\delta Q. \quad (6.3.17)$$

$$\begin{aligned} \Delta Q &= \int \delta Q \\ &= \int \delta W \\ &= \int_{V_1}^{V_2} p dV \\ &= \int_{V_1}^{V_2} \frac{RT}{V} dV \\ &= RT \ln \frac{V_2}{V_1} \end{aligned} \quad (6.3.18)$$

Definition 6.3.1.5 (Adiabatic changes). A change is **adiabatic** if it is both **thermally isolated** and **reversible**. This is also called an **isentropic** process.

Lemma 6.3.1.1 (Properties of adiabatic changes).

$$\mathrm{d}Q = 0 \quad (6.3.19a)$$

$$\mathrm{d}U = \mathrm{d}W \quad (6.3.19b)$$

For an **ideal gas**, therefore

$$C_V \mathrm{d}T = -p \mathrm{d}V = -\frac{RT}{V} \mathrm{d}V \quad (6.3.20a)$$

$$\ln \frac{T_2}{T_1} = -\frac{R}{C_V} \ln \frac{V_2}{V_1}. \quad (6.3.20b)$$

Definition 6.3.1.4 gives

$$\gamma = \frac{C_p}{C_V} = 1 + \frac{R}{C_V}, \quad (6.3.21)$$

Rearranging with Equation 6.3.20b we arrive at

Theorem 6.3.1.3 (Adiabatic equation).

$$TV^{\gamma-1} = \text{constant},$$

$$pV^{\gamma} = \text{constant},$$

$$T \left(\frac{T}{V} \right)^{\gamma-1} = \text{constant}, \quad (6.3.22)$$

$$p^{1-\gamma} T^{\gamma} = \text{constant}.$$

6.3.2 The Second Law

6.3.2.1 Heat Engines

The second law concerns itself with the **direction** of heat flow as the system approaches equilibrium.

Definition 6.3.2.1 (The Second Law). A statement of the second law is that no process is possible whose sole result is the transfer of heat from a colder to a hotter body. **OR** that no process is possible whose sole result is the complete conversion of heat into work.

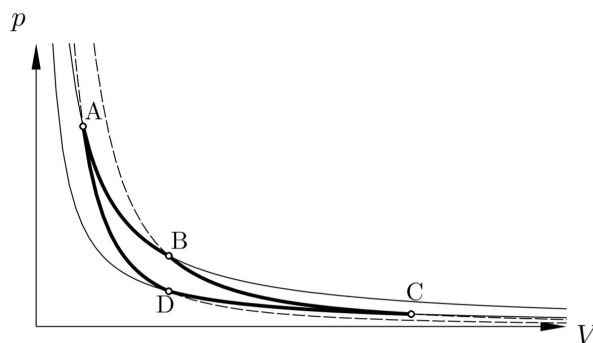


Figure 6.1: A p-V plot of the Carnot cycle.

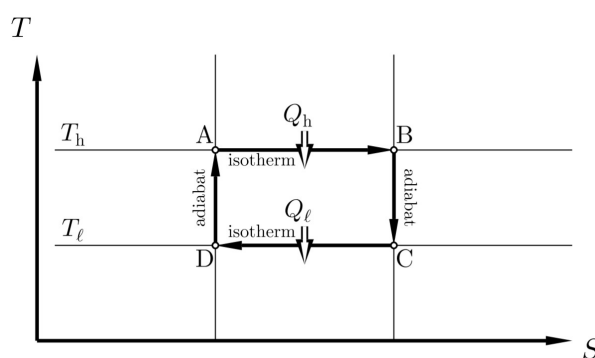


Figure 6.2: A T-S plot of the Carnot cycle. S here remains an undefined quantity that we assert is a function of pV^γ for now.

Definition 6.3.2.2 (Heat engine). An **engine** is a system operating a *cyclic* process that converts heat into work.

An example of a heat engine is the Carnot engine.

Definition 6.3.2.3 (Carnot engine). A **Carnot engine** runs on the Carnot cycle, consisting of two heat reservoirs of different temperatures, as shown in Figures 6.1 and 6.2. Thermodynamically, it consists of four alternating reversible isotherms and adiabats. Heat only enters and leaves on the isotherms whereas work is performed on all four segments.

Theorem 6.3.2.1 (Work done by Carnot engine). As this process is cyclic, we can conclude that

$$W = Q_h - Q_l \quad (6.3.23)$$

We can write down the governing equations for each segment:

$$A \rightarrow B : Q_h = RT_h \ln \frac{V_B}{V_A}, \quad (6.3.24a)$$

$$B \rightarrow C : \left(\frac{T_h}{T_l} \right) = \left(\frac{V_C}{V_B} \right)^{\gamma-1}, \quad (6.3.24b)$$

$$C \rightarrow D : Q_l = -RT_l \ln \frac{V_D}{V_C}, \quad (6.3.24c)$$

$$D \rightarrow A : \left(\frac{T_l}{T_h} \right) = \left(\frac{V_A}{V_D} \right)^{\gamma-1}. \quad (6.3.24d)$$

Equations 6.3.24b and 6.3.24d lead to

$$\frac{V_B}{V_A} = \frac{V_C}{V_D}, \quad (6.3.25)$$

and dividing Equation 6.3.24a by Equation 6.3.24c and substituting in Equation 6.3.25 gives

$$\frac{Q_h}{Q_l} = \frac{T_h}{T_l} \quad (6.3.26)$$

Definition 6.3.2.4 (Efficiency).

$$\eta = \frac{W}{Q_h} \quad (6.3.27)$$

Lemma 6.3.2.1 (Efficiency of Carnot engine).

$$\begin{aligned} \eta_{Carnot} &= \frac{Q_h - Q_l}{Q_h}, \\ \eta_{Carnot} &= \frac{T_h - T_l}{T_h}, \\ &= 1 - \frac{T_l}{T_h}. \end{aligned} \quad (6.3.28)$$

Theorem 6.3.2.2 (Carnot's theorem). Carnot's Theorem states that, of all the heat engines working between two given temperatures, none is more efficient than a Carnot engine.

Proof. We prove the theorem by contradiction. Suppose there exists engine E such that $\eta_E > \eta_{Carnot}$. Engine E and a Carnot engine running in reverse (having work done to

it) are connected together with two heat reservoirs at T_h and T_l respectively as shown in Figure 6.3. We have,

$$\begin{aligned}\frac{W}{Q'_h} &> \frac{W}{Q_h}, \\ Q_h &> Q'_h.\end{aligned}\tag{6.3.29}$$

We also have,

$$\begin{aligned}W &= Q'_h - Q'_l = Q_h - Q_l, \\ Q_h - Q'_h &= Q_l - Q'_l.\end{aligned}\tag{6.3.30}$$

Now Equation 6.3.30 means this compound engine is extracting a positive (due to Equation 6.3.29) amount of energy from the low temperature reservoir and dumping the same amount of energy into the high temperature reservoir. This is in direct violation of Definition 6.3.2.1 and therefore E cannot exist. \square

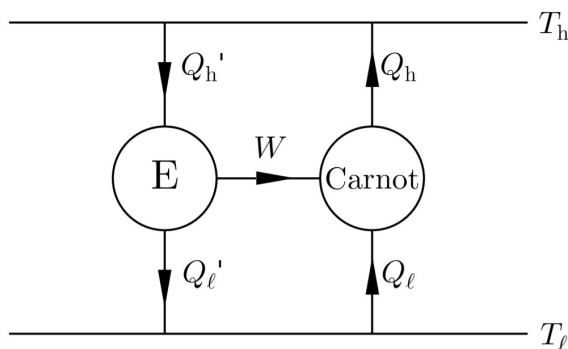


Figure 6.3: Proof of optimality of Carnot engine.

Corollary 6.3.2.1. All reversible engines working between two temperatures have the same efficiency η_{Carnot}

Proof. Suppose there is another reversible engine R such that $\eta_R \leq \eta_{Carnot}$. We run it in reverse and connect to a forward Carnot engine. This arrangement will again transfer heat from the cold reservoir to the hot one, *unless* they have the same efficiency. \square

Proposition 6.3.2.1 (Equivalence of Clausius and Kelvin statements). If a system violates Kelvin's statement of the second law, it also violates Clausius's statement, and *vice versa*.

Proof. As shown in Figure 6.4, by connecting a Kelvin violator, which extracts Q'_h and outputs the same amount of W , to a Carnot engine, which extracts Q_l , and outputs Q_h , we arrive at the same conclusion that this setup violates Clausius statement.

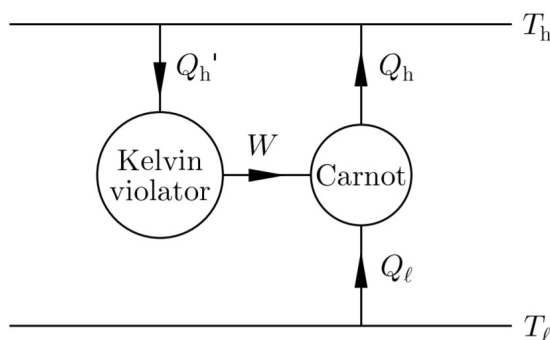


Figure 6.4

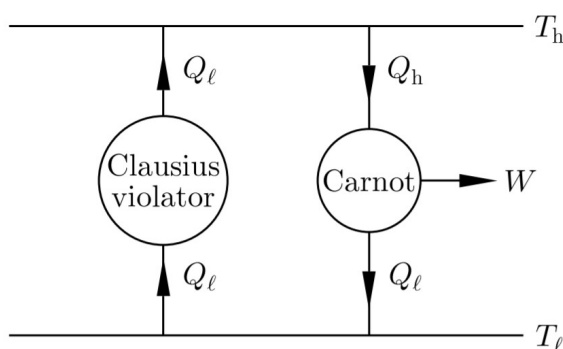


Figure 6.5

Now, in Figure 6.5 we connect a Clausius violator, transferring Q_l from cold to hot reservoir, and a Carnot engine, operating forward. The net action of the setup is to convert $Q_h - Q_l$ completely into work, hence violating Kelvin's statement. \square

6.3.2.2 Clausius' theorem

Consider a Carnot cycle. Heat is *not* conserved around a cycle, however Equation 6.3.26 states the important result that

$$\frac{Q_h}{Q_l} = \frac{T_h}{T_l}, \quad (6.3.31)$$

and so we are inspired to define ΔQ_{rev} as the heat entering the system at each segment, and therefore we have

$$\sum_{cycle} \frac{\Delta Q_{rev}}{T} = \frac{Q_h}{T_h} + \frac{(-Q_l)}{T_l} = 0, \quad (6.3.32)$$

which means $\Delta Q_{rev}/T$ sums to zero around the cycle. Replacing the sum by an integral we could write

$$\oint \frac{\Delta Q_{rev}}{T} = 0 \quad (6.3.33)$$

for the Carnot cycle.

To make a generalised case, we consider a general thermodynamic cycle in Figure 6.6(a), where heat δQ_i enters (or leaves, depends on the sign taken up) at a point in the cycle. At this point the system is connected to a reservoir at T_i . The total work extracted from the cycle is

$$\Delta W = \sum_{\text{cycle}} \delta Q_i. \quad (6.3.34)$$

Now suppose, as in Figure 6.6(b) that all the δQ_i is provided by Carnot engine(s) C_i ,

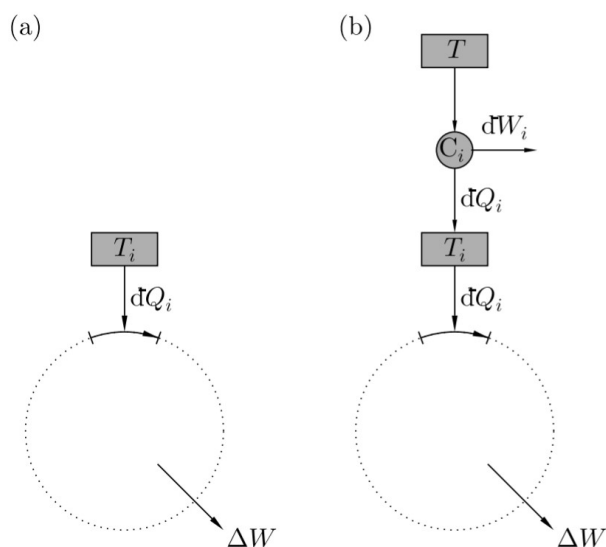


Figure 6.6

whose input reservoir is fixed at T and output reservoir is the variable T_i . The Carnot engine itself outputs work δW_i . We know, for a Carnot engine,

$$\begin{aligned} \frac{\text{heat to reservoir at } T_i}{T_i} &= \frac{\text{heat from reservoir at } T}{T} \\ \frac{\delta Q_i}{T_i} &= \frac{\delta Q_i + \delta W_i}{T} \\ \delta W_i &= \delta Q_i \left(\frac{T}{T_i} - 1 \right). \end{aligned} \quad (6.3.35)$$

For the system to not violate Kelvin's statement of second law, *i.e.*, not to convert all the

heat into work, as there is no output heat, we must require that

$$\begin{aligned}
 \text{output work} &\leq 0 \\
 \sum_{\text{cycle}} \mathrm{d}W_i + \Delta W &\leq 0 \\
 T \sum_{\text{cycle}} \frac{\mathrm{d}Q_i}{T_i} &\leq 0 \\
 \oint \frac{\mathrm{d}Q}{T} &\leq 0,
 \end{aligned} \tag{6.3.36}$$

which is the **Clausius inequality**.

Theorem 6.3.2.3 (Clausius's theorem). For any closed cycle, $\oint \frac{\mathrm{d}Q}{T} \leq 0$, where equality necessarily holds for a reversible cycle.

6.3.3 Entropy

We now are in position to introduce the thermodynamic quantity of entropy. According to Theorem 6.3.2.3, the integral

$$\int_A^B \frac{\mathrm{d}Q_{\text{rev}}}{T} \tag{6.3.37}$$

is path independent. Therefore the quantity $\mathrm{d}Q_{\text{rev}}/T$ is an *exact* differential.

Definition 6.3.3.1 (Entropy). Entropy is defined as a state function

$$\mathrm{d}S = \frac{\mathrm{d}Q_{\text{rev}}}{T}. \tag{6.3.38}$$

Proposition 6.3.3.1 (Adiabatic entropy change). For an adiabatic process (*defined as an adiathermal reversible process*), we have

$$\mathrm{d}Q_{\text{rev}} = 0. \tag{6.3.39}$$

Hence an adiabatic process involves no change in entropy, *i.e.*, isentropic.

Theorem 6.3.3.1 (Entropy inequality). Consider a cycle with a reversible and an irreversible segment, Theorem 6.3.2.3 gives that

$$\oint \frac{\mathrm{d}Q}{T} \leq 0$$

$$\int_A^B \frac{\mathrm{d}Q}{T} + \int_B^A \frac{\mathrm{d}Q_{rev}}{T} \leq 0 \quad (6.3.40)$$

$$\int_A^B \frac{\mathrm{d}Q}{T} \leq \int_A^B \frac{\mathrm{d}Q_{rev}}{T}$$

$$\mathrm{d}S = \frac{\mathrm{d}Q_{rev}}{T} \geq \frac{\mathrm{d}Q}{T}.$$

If we consider a thermally isolated system, *i.e.*, $\mathrm{d}Q = 0$, we then have

$$\mathrm{d}S \geq 0. \quad (6.3.41)$$

Now we are in a position to revisit the first law:

Theorem 6.3.3.2 (The first law). For a reversible change only, we have that

$$\mathrm{d}Q = T \mathrm{d}S \quad (6.3.42a)$$

$$\mathrm{d}W = -p \mathrm{d}V. \quad (6.3.42b)$$

For an irreversible change $\mathrm{d}Q \leq T \mathrm{d}S$ and $\mathrm{d}Q \geq -p \mathrm{d}V$, but we *always* have that

$$\mathrm{d}U = T \mathrm{d}S - p \mathrm{d}V. \quad (6.3.43)$$

We observe that S and V are the **natural variables** of U , and that they are both *extensive*. p and T however are *intensive* and behave somewhat like *forces*

6.4 MATHEMATICAL METHODS

6.4.1 Legendre transforms

This subsection is adapted from [22, 25] and Appendix F of [8].

Legendre transform, as we will see, is really just a long way to say ‘In a right-angled triangle, the slope (tangent) times the adjacent side equals the opposite side’.

For a function $F(x)$, we have all the information of the function F stored as ordered pairs of values $(x_i, F(x_i))$. There are many ways to store the same information of the function, one of which is through the gradient. For a function F whose second derivative is always positive, *i.e.*, F is monotonously increasing, the gradient $s(x) = dF(x)/dx$ is one-one to x , therefore we can invert it to get a single-valued function $x(s)$.

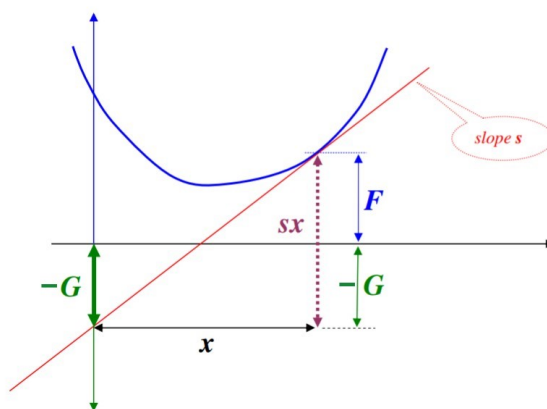


Figure 6.7: Sometimes the opposite sign convention for G is used to confer more symmetry.

Definition 6.4.1.1 (Legendre transform). As shown in Figure 6.7, we are inspired to write

$$sx = F + (-G). \quad (6.4.1)$$

Rearranged and showing explicitly the functional relationships,

$$G(s) = F(x(s)) - sx(s), \quad (6.4.2)$$

with $-G(s)$ (note the negative sign) known as the Legendre transform of $F(x)$. This is, at the root of it, just the geometrical statement at the beginning of the section.

Property 6.4.1.1 (Inverse). The inverse of the Legendre transform is the starting function, *i.e.*, the Legendre transform is its own inverse. We can see this by starting with

$$y(s) = -\frac{dG}{ds} \quad (6.4.3)$$

and inverting the monotonic function $y(s)$ to $s(y)$. We then construct the Legendre transform, $H(y)$ of $G(s)$:

$$\begin{aligned} H(y) &= G(s(y)) - ys(y) \\ G &= H - sy \end{aligned} \quad (6.4.4)$$

Comparison with Equation 6.4.2 reveals that we can identify $\{H, y\}$ with $\{F, x\}$. This gives us a recursive relation

$$s(x) = \frac{dF}{dx} \quad (6.4.5a)$$

$$x(s) = -\frac{dG}{ds}. \quad (6.4.5b)$$

Equation 6.4.5b can actually be obtained by differentiating Equation 6.4.2:

$$\frac{dG(s)}{ds} = \frac{dF(x(s))}{ds} - x(s) - s \frac{dx}{ds} = -x(s) - s \frac{dx}{ds} + \frac{dF}{dx} \frac{dx}{ds} = -x(s). \quad (6.4.6)$$

The symmetric is perhaps best displayed as

$$-G(s) + F(x) = sx. \quad (6.4.7)$$

Make no mistake here - there is only one independent variable: *either* x *or* s , with the other variable written as a function of it. The pair of variables x and s are called *conjugate* variables.

Property 6.4.1.2 (Extrema). Remembering that F is defined upwards positive and G downwards positive, as shown in Figure 6.7, we have

$$F_{\min} = -G(0) \quad (6.4.8a)$$

$$-G_{\min} = -F(0). \quad (6.4.8b)$$

This can be expressed in a more symmetrical form, with the RHS in Equation 6.4.7 vanishing in either case:

$$-G(0) + F(x_{\text{ext}}) = 0 \quad \text{and} \quad -G(s_{\text{ext}}) + F(0) = 0. \quad (6.4.9)$$

We now look at functions of more than one variables, for example, $F(x, y)$. We require that F is convex in x for all y and vice versa. The differential is

$$dF = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy \equiv p dx + v dy, \quad (6.4.10)$$

where $\{p, x\}$ and $\{v, y\}$ are conjugate pairs of variables, and p and v can be understood as the gradient in the two basis vector directions.

The Legendre transform aims, again, to swap one variable for its conjugate variable, say x for p . So essentially we ‘flatten’ a multivariate function into one dimension and the Legendre transform is

$$G(p, y) \equiv F(x(p), y) - px(p). \quad (6.4.11)$$

We see in the differential form that indeed the independent variable x is switched into p :

$$dG = dF - p dx - x dp = -x dp + v dy. \quad (6.4.12)$$

We can identify

$$x = -\frac{\partial G}{\partial p} \quad (6.4.13a)$$

$$v = \frac{\partial G}{\partial y}. \quad (6.4.13b)$$

Worked example 6.4.1.1. The internal energy is

$$U = U(S, V, \mathbf{N}), \quad (6.4.14)$$

with total differential given by

$$dU = T dS - P dV + \sum \mu_i dN_i \quad (6.4.15)$$

We know from Equation 6.2.3b that $-p$ (again note the negative sign) and V are *defined* by the fundamental thermodynamic equations to be conjugate variables. We therefore wish to define a new quantity $H = H(S, p, \mathbf{N})$ via a Legendre transform:

$$H(S, p, \mathbf{N}) \equiv U + pV. \quad (6.4.16)$$

We can immediately recover from the properties of Legendre transforms that

$$V = \frac{\partial H}{\partial p} \quad (6.4.17a)$$

$$T = \frac{\partial H}{\partial S} \quad (6.4.17b)$$

$$\mu_i = \frac{\partial H}{\partial N_i}, \quad (6.4.17c)$$

where subscripts are omitted for clarity.

The function H being a Legendre transform of U , encodes the exact same information as U but are easier to work with in situations where pressure is constant, as the total differential simplifies.

6.4.2 Euler-Maclaurin formula

This is adapted from [6].

We wish to approximate a sum

$$S(n) = \sum_{k=0}^n f(k) \quad (6.4.18)$$

by an integral of the form

$$\sigma(n) = \int_0^n f(x) dx. \quad (6.4.19)$$

But this is inaccurate for fast changing sums and it'll be more accurate to approximate behaviour of the function near each integer value, which is to say that we write the integral this way

$$\sigma(n) = \int_0^1 [f(0+x) + f(1+x) + f(2+x) + \cdots + f(n-1+x)] dx. \quad (6.4.20)$$

We approximate the behaviour of the function about each integer with a Taylor series expansion:

$$f(k+x) = \sum_{j=0}^{\infty} \frac{f^{(j)}(k)}{j!} x^j, \quad (6.4.21)$$

whose integral from 0 to 1 is

$$\int_0^1 f(k+x) dx = \sum_{j=0}^{\infty} \frac{f^{(j)}(k)}{(j+1)!} \quad (6.4.22)$$

So,

$$\sigma(n) = \sum_{m=0}^{\infty} \frac{S^{(m)}}{(m+1)!}. \quad (6.4.23)$$

Taking derivative of Equation 6.4.23 we have

$$\sigma^{(j)}(n) \quad (6.4.24)$$

6.4.3 Homogeneous functions

A polynomial

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \quad (6.4.25)$$

is of degree n . A polynomial is said to be homogeneous of degree n if all its terms are of the same degree n , for example,

$$x^2 + 5xy + 13y^2 \quad (6.4.26)$$

is homogeneous of degree 2. The same idea can be extended to functions: if for arbitrary λ

$$f(\lambda x) = \lambda^n f(x) \quad (6.4.27)$$

f is said to be homogeneous of degree n in the variable x , and

Definition 6.4.3.1 (Intensive and extensive variables). Intensive functions are homogeneous of degree zero and extensive functions are homogeneous of degree one.

Now consider

$$S = S(U, V, n) \quad (6.4.28)$$

physically, entropy is extensive, so we can write

$$S(\lambda U, \lambda V, \lambda n) = \lambda S(U, V, n) \quad (6.4.29)$$

If we choose $\lambda = 1/n$ we can write

$$S\left(\frac{U}{n}, \frac{V}{n}, 1\right) \equiv S_m(U_m, V_m) = \frac{1}{n} S(U, V, n) \quad (6.4.30)$$

so

$$nS_m(U_m, V_m) = S(U, V, n) \quad (6.4.31)$$

This illustrates why λ is known as the *scaling function*.

6.5 FREE ENERGIES

This section is adapted from [8].

In laboratory conditions, microscopic, extensive quantities such as S , U and N is difficult to keep track of at boundaries of systems. Instead we can control and keep track of macroscopic, intensive quantities such as p and T . This makes the use of $U(S, V, N)$ or $S(U, V, N)$ rather cumbersome. However, the second law still applies and as such we can devise new extremum principles to find conditions of equilibrium.

6.5.1 Helmholtz free energy

6.5.1.1 Inspiration

Imagine a test tube in a water bath at a fixed temperature T , which act as a thermal reservoir. Therefore we need to find a function A whose natural variables are (T, V, N) for us to construct a new extremum principle.

Now assume that the combined heat bath-test tube system isolated from its surroundings, equilibrium is predicted by the state of maximum entropy $S(U, V, N)$. Any change towards equilibrium must have that

$$dS_{\text{combined system}} = dS_{\text{bath}} + dS_{\text{test tube}} \geq 0. \quad (6.5.1)$$

Since the combined system is isolated,

$$dU_{\text{bath}} + dU_{\text{test tube}} = 0. \quad (6.5.2)$$

The fundamental equation gives

$$dS_{\text{bath}} = \left(\frac{1}{T}\right) dU + \left(\frac{p}{T}\right) dV - \left(\frac{\mu}{T}\right) dN = \left(\frac{1}{T}\right) dU_{\text{bath}} = -\frac{dU_{\text{test tube}}}{T}, \quad (6.5.3)$$

where the second equality comes from Equation 6.5.2. Applying Equation 6.5.1 further results in

$$dS_{\text{test tube}} - \frac{dU_{\text{test tube}}}{T} \geq 0 \quad (6.5.4a)$$

$$dU_{\text{test tube}} - T dS_{\text{test tube}} \leq 0, \quad (6.5.4b)$$

where we have successfully obtained a relation only in terms of the test tube quantities. We therefore define

Definition 6.5.1.1 (Helmholtz free energy).

$$A \equiv U - TS, \quad (6.5.5)$$

with differential

$$dA = dU - T dS - S dT = dU - T dS \quad \text{at constant temperature.} \quad (6.5.6)$$

Therefore we see that Helmholtz free energy is essentially the amount of reversible work obtainable from an isothermal and isochoric system:

$$\begin{aligned} dA &= \delta Q + \delta W - T dS \\ &= \delta W, \end{aligned} \quad (6.5.7)$$

where we assumed reversibility, which extracts most amount of work.

Property 6.5.1.1 (Condition of equilibrium). We can see that the condition in Equation 6.5.1 is fulfilled when

$$(dA)_T \leq 0. \quad (6.5.8)$$

This means that equilibrium is reached when A is at a minimum.

Property 6.5.1.2 (Fundamental equation).

$$dA = -S dT - p dV + \sum_{j=1}^M \mu_j dN_j \quad (6.5.9)$$

And we can trivially obtain more definitions:

$$S = - \left(\frac{\partial F}{\partial T} \right)_{V, N} \quad (6.5.10a)$$

$$p = - \left(\frac{\partial F}{\partial V} \right)_{T, N} \quad (6.5.10b)$$

$$\mu_j = \left(\frac{\partial A}{\partial N_j} \right)_{V, T, N_{i \neq j}}. \quad (6.5.10c)$$

On the surface this definition seems strange because we need to hold T and V constant and we're only left with the chemical potential term. It is true that if we're looking at a one-component non-reacting system, if we fix T and V , no change can take place. But in a multi-component, reacting mixture with fixed T and V we have

$$dA = \sum_{j=1}^M \mu_j dN_j \quad (6.5.11)$$

where A can depend on T and V in the *equation of state*, we have the chemical potential terms in the differentials, which changes to minimise the entropy. The following two examples illustrates this idea.

6.5.1.2 A microscopic look at Helmholtz I: Dimerisation

Suppose $N = 2$ gas particles are contained in a test tube of V lattice sites in a row at T . We ask under what conditions will the two associate into a dimer. We compute the free energies of monomer and dimer states and compare them.

Dimers: Suppose when two monomer sit on adjacent sites a 'bond energy' $U = -\epsilon$ is additionally introduced.

Within V sites, multiplicity for a dimer is

$$W_d = V - 1 \approx V. \quad (6.5.12)$$

The Helmholtz free energy is then given as

$$A_d = U_d - TS_d = -\epsilon - kT \ln V. \quad (6.5.13)$$

Monomers: There is no energy of interaction, and the multiplicity is

$$W_m = W_{\text{total}} - W_d = \frac{V!}{(2!)(V-2)!} - (V-1) = \left(\frac{V}{2} - 1\right)(V-1) \approx \frac{V^2}{2}. \quad (6.5.14)$$

So the free energy is

$$A_m = -kT \ln \left(\frac{V^2}{2} \right). \quad (6.5.15)$$

Temperature at which both are equally stable is, after some algebra

$$T_0 = \frac{\epsilon}{k \ln(V/2)}. \quad (6.5.16)$$

Had we only maximised entropy we would conclude that the dimers will never even form, hence missing out on the significant role energy plays at lower temperatures here.

6.5.1.3 A microscopic look at Helmholtz II: Polymer collapse

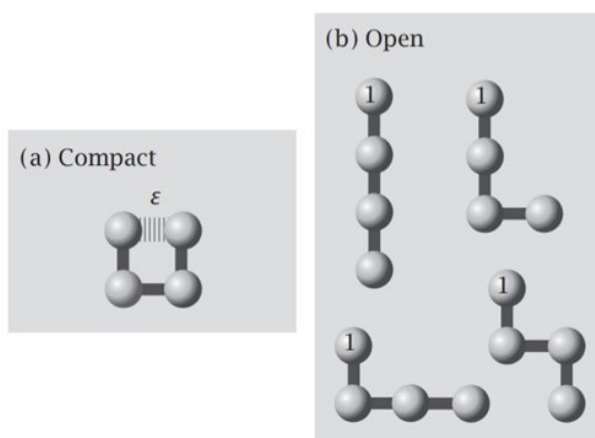


Figure 6.8: The two energy levels.

Consider a two-dimensional model polymer with four monomers in a closed test tube solution in a bath at T and the ends of the polymer are attracted to each other by $U = -\epsilon$. The multiplicity is 1 as only one conformation can result in the close-ring polymer. The open-ring state has 4 other multiplicities. We compute the free energies:

$$A_c = -\epsilon - kT \ln 1 = -\epsilon \quad (6.5.17a)$$

$$A_o = -kT \ln 4. \quad (6.5.17b)$$

The collapse temperature T_0 is given as

$$T_0 = \frac{\epsilon}{k \ln 4}. \quad (6.5.17c)$$

We note that for this process,

$$\Delta F_{\text{collapse}} = -\epsilon + kT \ln 4 \quad (6.5.18a)$$

$$\Delta U_{\text{collapse}} = -\epsilon \quad (6.5.18b)$$

$$\Delta S_{\text{collapse}} = -k \ln 4 \quad (6.5.18c)$$

As no volume change is involved, in a reversible process no work is done and $\Delta U = -\epsilon$ and the process is exothermic at all temperature.

6.5.2 Enthalpy

The enthalpy is most commonly obtained as a Legendre transform of U as explained in Worked example 6.4.1.1. We can write out the total differential:

$$dH = T dS + V dP + \sum \mu_j dN_j. \quad (6.5.19)$$

At constant pressure, with a non-reacting mixture, we can write

$$dS = \frac{dH}{T}. \quad (6.5.20)$$

By way of the argument used in Section 6.5.1.1, we introduce two systems A and B , with A being an open test tube and B the surroundings, acting as a pressure reservoir. Any change towards equilibrium will have

$$dS_{A+B} = dS_A + dS_B \geq 0. \quad (6.5.21)$$

Because B is also a temperature reservoir, we write

$$dS_B = -\frac{dQ_A}{T}. \quad (6.5.22)$$

Combined, we have

$$dS_A - \frac{dQ_A}{T} \geq 0. \quad (6.5.23)$$

Under constant pressure we have

$$dH = dU + p dV - V dp = \delta Q - p dV + p dV - V dp = \delta Q, \quad (6.5.24)$$

and as such

$$T dS_A - dH \geq 0. \quad (6.5.25)$$

For tending to equilibrium inside A , we require that

$$dS_A \geq 0, \quad (6.5.26)$$

therefore

$$\frac{\delta Q_A}{T} \leq 0. \quad (6.5.27)$$

So finally we have that

$$dH \leq 0 \quad (6.5.28)$$

for processes to tend to equilibrium at constant pressure.

6.5.3 Gibbs free energy

To get a function that is minimised at equilibrium at constant pressure and temperature, we perform a Legendre transform on $H(S, p, \mathbf{N})$ to get $G(T, p, \mathbf{N})$, noting that T and S are a conjugate pair and the usual sign convention applies:

$$G(T, p, \mathbf{N}) = H - TS. \quad (6.5.29)$$

The total differential is given as

$$dG = -S dT + V dp + \sum \mu_j dN_j. \quad (6.5.30)$$

At constant pressure and temperature,

$$dG = dH - T dS = \delta Q - T dS. \quad (6.5.31)$$

Invoking Clausius theorem we can show the equilibrium condition, but to do so explicitly, we use the second law, for

$$dS_{A+B} = dS_A + dS_B = dS_A - \frac{\delta Q_A}{T} \geq 0, \quad (6.5.32)$$

under constant pressure and temperature, this becomes

$$\begin{aligned} T \, dS_A - dH_A &\geq 0 \\ (dG)_{p,T} &\leq 0. \end{aligned} \quad (6.5.33)$$

Therefore G is minimised at equilibrium.

It is called ‘free energy’ because it shows how much non-mechanical work can be extracted, which is important in chemistry and biology, again we assume constant pressure and temperature:

$$dG = dU + p \, dV - T \, dS \quad (6.5.34a)$$

$$= T \, dS - p \, dV + dW_x + p \, dV - T \, dS \quad (6.5.34b)$$

$$= dW_x, \quad (6.5.34c)$$

where we assume the process is reversible, and the x subscript means other forms of work like electrical work.

6.5.4 More on thermodynamic functions

6.5.4.1 Functional dependencies

Below we give a summary of the fundamental functions.

| Function | Extremum at eqm | Fundamental equation | Definition | Useful for |
|--------------|-----------------|---|---------------------------------|---------------------------------|
| $U(S, V, N)$ | Minimum | $dU = T \, dS - p \, dV + \sum \mu_j \, dN_j$ | | |
| $S(U, V, N)$ | Maximum | $dS = \left(\frac{1}{T}\right) dU + \left(\frac{p}{T}\right) dV + \sum \left(\frac{\mu_j}{T}\right) dN_j$ | | |
| $H(S, p, N)$ | Minimum | $dH = T \, dS + V \, dp + \sum \mu_j \, dN_j$ | $H \equiv U + pV$ | Const S and p , $dH = dU$ |
| $A(T, V, N)$ | Minimum | $dA = -S \, dT - p \, dV + \sum \mu_j \, dN_j$ | $A \equiv U - TS$ | Const T and V , $dA = dW$ |
| $G(T, p, N)$ | Minimum | $dG = -S \, dT + V \, dp + \sum \mu_j \, dN_j$ | $G \equiv H - TS \equiv A + pV$ | Const T and p , $dG = dW_x$ |

We should be aware that, for example the function U in $A = U - TS$ is not $U(S, V, N)$ but instead takes on the arguments of A such that $U = U(T, V, N)$. This is why minimising $U(T, V, N)$ or maximising $S(T, V, N)$ individually does not give us equilibrium but only their sum in this specific functional form, when minimised, gives equilibrium.

We illustrate this idea by considering

$$G(T, p, N) = H(T, p, N) - TS(T, p, N). \quad (6.5.35)$$

H and S here are non-fundamental functions, but they are useful in that they can be conveniently measured and calculated. Now because at constant pressure,

$$C_p = \left(\frac{dQ}{dT} \right)_p = \left(\frac{\partial H}{\partial T} \right)_p = T \left(\frac{\partial S}{\partial T} \right)_p, \quad (6.5.36)$$

we can write that

$$\Delta H(T, p) = \int_{T_A}^{T_B} C_p(T) dT \quad (6.5.37a)$$

$$\Delta S(T, p) = \int_{T_A}^{T_B} \frac{C_p(T)}{T} dT. \quad (6.5.37b)$$

From a series of constant-pressure heat capacity experiments one can obtain $G(T, p)$.

From constant-volume heat capacity experiments one can similarly obtain $A(T, V) = U(T, V) - TS(T, V)$ since we know that at constant volume,

$$C_V = \left(\frac{\partial U}{\partial T} \right)_V = T \left(\frac{\partial S}{\partial T} \right)_V, \quad (6.5.38)$$

we can write

$$\Delta U(T, V) = \int_{T_A}^{T_B} C_V(T) dT \quad (6.5.39a)$$

$$\Delta S(T, V) = \int_{T_A}^{T_B} \frac{C_V(T)}{T} dT. \quad (6.5.39b)$$

6.5.4.2 Non-standard states

We can further write, for ideal mixtures,

$$\begin{aligned} dG &= \frac{nRT}{p} dp \\ \int_{p_1}^{p_2} dG &= \int_{p_1}^{p_2} \frac{nRT}{p} dp \\ G(p_2) - G(p_1) &= nRT \ln \frac{p_2}{p_1} \\ G(p) &= G^\circ + nRT \ln \frac{p}{p^\circ} \end{aligned} \quad (6.5.40)$$

For ideal gases, the chemical potential is the molar Gibbs energy, so we can always write

$$\mu_i(p_i) = \mu_i^\circ + RT \ln \frac{p_i}{p^\circ} \quad (6.5.41)$$

and a parametric ($\mu(p) \rightarrow \mu(c(p))$) substitution gives

$$\mu_i(c_i) = \mu_i^\circ + RT \ln \frac{c_i}{c^\circ} \quad (6.5.42)$$

And for the generic reaction



with some algebra we get

$$\Delta_r G^\circ = -RT \ln K_p \quad (6.5.44)$$

and also

$$\Delta_r G^\circ = -RT \ln K_c \quad (6.5.45)$$

6.5.4.3 Gibbs-Helmholtz equation

For the function G/T , we can always write, with chain rule:

$$\frac{\partial}{\partial T}(GT^{-1}) = T^{-1} \frac{\partial G}{\partial T} - T^{-2} G \quad (6.5.46)$$

noting that $\partial G / \partial T = -S$ by definition, we proceed:

$$\begin{aligned} \frac{\partial}{\partial T}(GT^{-1}) &= -ST^{-1} - T^{-2}(H - TS) \\ &= -\frac{H}{T^2} \end{aligned} \quad (6.5.47)$$

where we have derived the

Theorem 6.5.4.1 (Gibbs-Helmholtz equation). The equation states that

$$\frac{\partial}{\partial T} \left(\frac{G}{T} \right) = -\frac{H}{T^2} \quad (6.5.48)$$

From this we can derive the van't Hoff equation

Theorem 6.5.4.2 (van't Hoff equation). The equation states that

$$\frac{\partial \ln K}{\partial T} = \frac{\Delta_r H^\circ}{RT^2} \quad (6.5.49)$$

Proof. The free energies in Gibbs-Helmholtz equation can be replaced by changes (over

a reaction):

$$\frac{\partial}{\partial T} \left(\frac{\Delta_r G^\circ}{T} \right) = -\frac{\Delta_r H^\circ}{T^2} \quad (6.5.50)$$

and invoking eqs. 6.5.44 and 6.5.45 we can write that

$$\frac{\partial}{\partial T} (-R \ln K) = -\frac{\Delta_r H^\circ}{T^2} \quad (6.5.51)$$

which rearranges to give the van't Hoff equation. □

6.5.5 The Euler and Gibbs-Duhem equations

6.5.5.1 The Euler equation

Referring to section 6.4.3, we now consider the internal energy, which again is physically an extensive function so we can write¹

$$U(\lambda S, \lambda V, \lambda n) = \lambda U(S, V, n) \quad (6.5.52)$$

using the chain rule, we can write

$$\begin{aligned} U(S, V, n) &= \left(\frac{\partial U}{\partial(\lambda S)} \right)_{V,n} \frac{\partial(\lambda S)}{\partial \lambda} + \left(\frac{\partial U}{\partial(\lambda V)} \right)_{S,n} \frac{\partial(\lambda V)}{\partial \lambda} + \left(\frac{\partial U}{\partial(\lambda n)} \right)_{S,V} \frac{\partial(\lambda n)}{\partial \lambda} \\ &= \left(\frac{\partial U}{\partial(\lambda S)} \right)_{V,n} S + \left(\frac{\partial U}{\partial(\lambda V)} \right)_{S,n} V + \left(\frac{\partial U}{\partial(\lambda n)} \right)_{S,V} n \\ &= TS - pV + \mu n \end{aligned} \quad (6.5.53)$$

This is known as the **Euler equation**, which relates all seven thermodynamic variables.

We can further extend this to the Gibbs free energy, which can be written as

$$\begin{aligned} G &= H - TS = U + pV - TS \\ &= TS - pV + \mu n + pV - TS \\ &= \mu n \end{aligned} \quad (6.5.54)$$

Which can be simply extended to

$$G = \sum_i \mu_i N_i \quad (6.5.55)$$

¹We use n here for simplicity but this can be easily converted and extended to N

This holds true for all homogeneous systems, which can be non-ideal.

6.5.5.2 The Gibbs-Duhem equation

The Euler equation(s) might look strange because the total differential has extra terms that are not present in the master equation, and vice versa. We need to realise that the master equation started from the conservation of internal energy, which tells us that $U(S, V, N)$ must be a state function, and we can take the total differential and make the definitions to get

$$dU = T dS - p dV + \mu dN \quad (6.5.56)$$

The Euler equation on the other hand is a statement of the extensivity of energy. So both expressions are logical and the only way to reconcile the two is to take the extra terms in the differential form of the Euler equation to be zero:

$$\begin{aligned} dU &= T dS + S dT - p dV - V dp + \mu dN + N d\mu \\ &= \underbrace{T dS - p dV + \mu dN}_{dU} + \underbrace{S dT - V dp + N d\mu}_{=0} \end{aligned} \quad (6.5.57)$$

This gives us

Theorem 6.5.5.1 (The Gibbs-Duhem relation). The relation states that

$$S dT - V dp + \sum_i N_i d\mu_i = 0 \quad (6.5.58)$$

Which tells us that the three intensive variables are interrelated, and specifying two variables allows us to compute the third.

This can be extended to Gibbs free energy, where $G = \sum_i \mu_i N_i$ and the master equation is $dG = -S dT + V dp + \sum_i \mu_i dN_i$, and we can conclude

$$S dT - V dp + \sum_i N_i d\mu_i = 0 \quad (6.5.59)$$

Under constant T and p , we have

$$\sum_i N_i d\mu_i = 0 \quad (6.5.60)$$

which in the more illustrative case of a binary system, gives us

$$d\mu_B = -\frac{N_A}{N_B} d\mu_A \quad (6.5.61)$$

which shows that chemical potentials are functions of the composition N .

6.6 MAXWELL'S RELATIONS AND MIXTURES

Before we dive head in to proper statistical thermodynamics, let's push the classical, differential approach to its logical end.

There are many external constraints besides the likes of pressure and volume that we have been discussing, for example, a lipid bilayer can change its surface area, and a rubber band its length. We ask the following questions:

- What is the extra degree(s) of freedom, *i.e.*, new variables?
- What is the fundamental function and what does it depend on?
- What is the condition of equilibrium?
- Is the process tending to equilibrium driven by entropy or enthalpy or both?

6.6.1 Designing a fundamental equation

Definition 6.6.1.1 (Conjugate pairs). We have introduced the idea of conjugate pairs, but we now generalise the definition:

For a generalised extensive variable X_j , we have the conjugate force-analogue F_j that is an intensive variable given as

$$F_j = \left(\frac{\partial U}{\partial X_j} \right)_{S,V,\mathbf{N},X_{i \neq j}}. \quad (6.6.1)$$

Now to make the new, augmented internal energy function we just introduce the *intensive* variable into the argument to get

$$U = U(S, V, \mathbf{N}, \mathbf{X}) \quad (6.6.2a)$$

$$dU = T dS - p dV + \sum \mu_j dN_j + \sum F_j dX_j. \quad (6.6.2b)$$

Now, although the definition relies on internal energy, it is almost always easier to work with other thermodynamic potentials - because they are exactly introduced to make experiments easier. To transform to a suitable potential, we perform Legendre transforms on U .

Worked example 6.6.1.1. Suppose a system can change its surface area independently of its volume. What is the function that identifies the state of equilibrium?

Equation 6.6.1 tells us that

$$\gamma = \left(\frac{\partial U}{\partial A} \right)_{S,V,N}, \quad (6.6.3)$$

where γ is conjugate to surface area and is the *surface tension*. However this equation is not really useful as we have to control S and V , which, for example if we are investigating a lipid bilayer *in vivo*, we would want to control T and p instead. This is easy as we can just transform

$$U \leftrightarrow G = U + pV - TS, \quad (6.6.4)$$

and

$$dG = -S dT + V dp + \sum \mu_j dN_j + \gamma dA, \quad (6.6.5)$$

where

$$\gamma = \left(\frac{\partial G}{\partial A} \right)_{T,p,N}, \quad (6.6.6)$$

a quantity that is easy to determine experimentally. However the process can be further streamlined with Maxwell's relations.

6.6.2 Maxwell's relations

Maxwell's relations result from *equivalence of mixed partials*.

Suppose we want to find

$$\left(\frac{\partial S}{\partial p} \right)_{T,N}, \quad (6.6.7)$$

which could be useful for understanding how the entropies of materials change when they are squeezed. We identify the independent variables as T, p and N and identify that the natural function is $G(T, p, N)$, with total differential

$$dG = -S dT + V dp + \sum \mu_j dN_j. \quad (6.6.8)$$

To get the desired partial, we see that

$$S = -\frac{\partial G}{\partial T}, \quad (6.6.9)$$

so

$$-\frac{\partial S}{\partial p} = \frac{\partial^2 G}{\partial p \partial T} = \frac{\partial^2 G}{\partial T \partial p} = \frac{\partial V}{\partial T}. \quad (6.6.10)$$

Therefore we get our (unique) Maxwell's relation that

$$\left(\frac{\partial S}{\partial p} \right)_{T,N} = - \left(\frac{\partial V}{\partial T} \right)_{p,N}, \quad (6.6.11)$$

a quantity that is much easier to experimentally determine.

6.6.3 Susceptibilities

Definition 6.6.3.1 (Susceptibility). Susceptibility is generally defined as the fractional change in a quantity as a result of an infinitesimal change in another quantity:

$$\sigma = \frac{1}{X} \frac{\partial X}{\partial L}. \quad (6.6.12)$$

An example would be the thermal expansion coefficient α :

$$\alpha = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_p, \quad (6.6.13)$$

the fractional change in the volume of a system with temperature at constant pressure.

Coupled with Maxwell's relations, the experimentally easy to determine susceptibilities can usually give insight to hard to measure quantities. For example the thermal expansion coefficient α above, coupled with

$$\left(\frac{\partial S}{\partial p} \right)_{T,N} = - \left(\frac{\partial V}{\partial T} \right)_{p,N} \quad (6.6.14)$$

we can write

$$\alpha = - \frac{1}{V} \left(\frac{\partial S}{\partial p} \right)_{T,N}. \quad (6.6.15)$$

This means if $\alpha \geq 0$, increasing pressure orders the system. Further coupled with equations of state like the ideal gas law we can get

$$\alpha = \frac{p}{NkT} \frac{Nk}{p} = \frac{1}{T} \geq 0, \quad (6.6.16)$$

which says that increasing pressure decreases the entropy of an ideal gas.

Change in entropy can then be easily calculated at constant pressure:

$$dS = \left(\frac{\partial S}{\partial p} \right)_{T,N} dp = - \left(\frac{\partial V}{\partial T} \right)_{p,N} dp = -\alpha V dp \quad (6.6.17a)$$

$$\Rightarrow \Delta S = - \int_{p_1}^{p_2} \alpha(p) V(p) dp. \quad (6.6.17b)$$

Another example of susceptibility is

Definition 6.6.3.2 (Isothermal compressibility).

$$\kappa = -\frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_T. \quad (6.6.18)$$

A further example is about measuring $(\partial U/\partial V)_T$, which should tell us about the cohesive forces in materials. But this is hard to measure. We again recruit the help of susceptibilities. We write

$$dU = \left(\frac{\partial U}{\partial V} \right)_T dV + \left(\frac{\partial U}{\partial T} \right)_V dT \quad (6.6.19a)$$

$$dS = \left(\frac{\partial S}{\partial V} \right)_T dV + \left(\frac{\partial S}{\partial T} \right)_V dT, \quad (6.6.19b)$$

substituting Equation 6.6.19b into $dU = T dS - p dV$ we get

$$\left(\frac{\partial U}{\partial V} \right)_T dV + \left(\frac{\partial U}{\partial T} \right)_V dT = T \left[\left(\frac{\partial S}{\partial V} \right)_T dV + \left(\frac{\partial S}{\partial T} \right)_V dT \right] - p dV. \quad (6.6.20)$$

At constant T this reduces to

$$\left(\frac{\partial U}{\partial V} \right)_T = T \left(\frac{\partial S}{\partial V} \right)_T - p, \quad (6.6.21)$$

using the Maxwell relation

$$\left(\frac{\partial S}{\partial V} \right)_T = \left(\frac{\partial p}{\partial T} \right)_V \quad (6.6.22)$$

gives

$$\left(\frac{\partial U}{\partial V} \right)_T = T \left(\frac{\partial p}{\partial T} \right)_V - p. \quad (6.6.23)$$

The quantity $(\partial p/\partial T)_V$ is the *thermal pressure coefficient*. For typical liquids, $T(\partial p/\partial T)_V - p$ is negative at high densities and positive at low densities, which is exactly the behaviour we expect due to the dominance of repulsive forces at high densities and attractive forces at low densities.

6.6.4 Entropic and enthalpic components

We first look at a rubber band, whose length L , an extensive variable with conjugate re-stroing force f , is the new variable to be introduced as an argument of U while N is fixed:

$$dU = T dS - p dV + f dL. \quad (6.6.24)$$

Our experiment is to be carried out at constant T and p , so G is the natural function to work with. Under a Legendre transform we have

$$dG = d(U + pV - TS) = -S dT + V dp + f dL. \quad (6.6.25)$$

So we see

$$f = \left(\frac{\partial G}{\partial L} \right)_{T,p} = \left(\frac{\partial H}{\partial L} \right)_{T,p} - T \left(\frac{\partial S}{\partial L} \right)_{T,p}, \quad (6.6.26)$$

meaning the restoring force f is driven by a combination of entropy and enthalpy. To make the entropic component easier to measure, we find its Maxwell relationship:

$$\left(\frac{\partial S}{\partial L} \right)_{T,p} = - \frac{\partial^2 G}{\partial T \partial L} = - \left(\frac{\partial f}{\partial T} \right)_{p,L}, \quad (6.6.27)$$

and the enthalpic component is simply

$$\left(\frac{\partial H}{\partial L} \right)_{T,p} = f - T \left(\frac{\partial f}{\partial T} \right)_{p,L}, \quad (6.6.28)$$

both quantities are very easy to obtain from simple experiments.

Now we take a look at a Langmuir trough, a liquid container with a bar frictionlessly floating in it. On the right side is water and on the left side is water plus a surfactant like phospholipids. If a lateral pressure π is exerted towards the left, the surfactant surface must change its area, a . We identify the relevant function to be G and that

$$dG = -S dT + V dp - \pi dA, \quad (6.6.29)$$

where the negative sign in front of π means the free energy increases as the surfactant layer is compressed, rather like the pV sign convention. In an experiment we can determine $\pi(T, a, N)$, with N being the number of surfactant molecules.

Suppose N surfactant molecules can occupy A lattice sites on the left side, with a conversion factor λ of area per lattice site. In the dilute limit of $N \ll A$, the multiplicity is

$$W \approx A^N. \quad (6.6.30)$$

So

$$S(A) = k \ln W = Nk \ln A. \quad (6.6.31)$$

In Equation 6.6.29 we can partition π into the two components again:

$$\pi = - \left(\frac{\partial G}{\partial A} \right)_{T,p} = - \left(\frac{\partial H}{\partial A} \right)_{T,p} + T \left(\frac{\partial S}{\partial A} \right)_{T,p}, \quad (6.6.32)$$

and we can get a Maxwell relation

$$\left(\frac{\partial S}{\partial A}\right)_{T,p} = -\frac{\partial^2 G}{\partial T \partial A} = \left(\frac{\partial \pi}{\partial T}\right)_{p,A}. \quad (6.6.33)$$

From Equation 6.6.31, we have

$$\left(\frac{\partial \pi}{\partial T}\right)_{p,A} = \frac{Nk}{A}, \quad (6.6.34)$$

which implies that

$$\pi A = NkT, \quad (6.6.35)$$

the two dimensional ideal gas law. Comparing with Equation 6.6.32 we see that the pressure is purely entropic. However when the surfactant becomes more concentrated the molecules might start interacting amongst themselves and an enthalpic element may be seen.

6.6.5 Partial molar properties

We now look at multicomponent systems such as liquid mixtures and metal alloys.

Definition 6.6.5.1 (Partial molar volume).

$$v_j = \left(\frac{\partial V}{\partial n_j}\right)_{T,p,n_{i \neq j}}, \quad (6.6.36)$$

the change in volume when dn_j moles of j molecules are added to the system. Alternatively

$$dV = \sum_{j=1}^M \left(\frac{\partial V}{\partial n_j}\right)_{T,p,n_{i \neq j}} dn_j = \sum v_j dn_j \quad (6.6.37)$$

Chemical potentials can be variously defined

$$\begin{aligned} \mu_j &= \left(\frac{\partial U}{\partial N_j}\right)_{V,S,N_{i \neq j}} = \left(\frac{\partial G}{\partial N_j}\right)_{T,p,N_{i \neq j}} \\ &= \left(\frac{\partial A}{\partial N_j}\right)_{T,V,N_{i \neq j}} = \left(\frac{\partial H}{\partial N_j}\right)_{S,p,N_{i \neq j}}. \end{aligned} \quad (6.6.38)$$

However, the definition of partial molar quantities are defined specifically to be quantities measured at constant T and p , so

Definition 6.6.5.2 (Partial molar free energy).

$$\mu_{j,m} = \left(\frac{\partial G}{\partial n_j} \right)_{T,p,n_{i \neq j}} \quad (6.6.39)$$

We can divide the partial molar free energy into its partial molar enthalpy and entropy components:

$$\begin{aligned} \mu_j &= \left(\frac{\partial G}{\partial n_j} \right)_{T,p,n_{i \neq j}} \\ &= \left(\frac{\partial H}{\partial n_j} \right)_{T,p,n_{i \neq j}} - T \left(\frac{\partial S}{\partial n_j} \right)_{T,p,n_{i \neq j}} \\ &\equiv h_j - T s_j. \end{aligned} \quad (6.6.40)$$

Chapter 7

Statistical mechanics

7.1 THE BOLTZMANN DISTRIBUTION LAW

7.1.1 The canonical ensemble

The canonical ensemble consists of cells of constant V, T, N , essentially a massive collection of bubbles sitting in a heat bath, and asks the question of how many of the bubbles will have the same energy. It is important to note that the particles within each cells are allowed to interact. In the derivation below note how we never assume the particles to be independent (interacting).

7.1.1.1 Derivation

Consider an isolated supersystem with a bath (surroundings) and a system we're interested in. We have

$$E_{\text{tot}} = E_{\text{sys}} + E_{\text{bath}}. \quad (7.1.1)$$

We now consider two states of the system, **A**, with $E_{\text{bath,A}} = E_{\text{tot}}$ and **B**, with the system in state m with E_m . The number of microstates available to the supersystem Ω_{tot} , is constant. We can write, from the postulate of equal probabilities, we can write

$$P_m \propto \Omega_{\text{bath}}(E_{\text{tot}} - E_m) \quad (7.1.2)$$

This can look superficially simple but let's think what it actually means: the postulate of equal probabilities only applies to the *supersystem*, which is isolated. So we can only say

that the probability of observing the *system* with energy E_m is proportional to the number of microstates that has the bath with that energy. The choice to use the perspective of the bath but not the system is arbitrary now, however a subsequent approximation makes it necessary.

To find the exact functional form of $\Omega_{\text{bath}}(E_{\text{bath}})$, we invoke the definition

$$\left(\frac{\partial S}{\partial U}\right)_V = \frac{1}{T} \quad (7.1.3)$$

Considering just the bath, we have

$$\frac{d \ln \Omega_{\text{bath}}(E_{\text{bath}})}{dE_{\text{bath}}} = \frac{1}{kT}. \quad (7.1.4)$$

Now we make the important approximation that, in the limit of $E_{\text{tot}} \gg E_m$, we can approximate the derivative with a simple geometric gradient

$$\frac{\ln \Omega_{\text{bath}}(E_{\text{bath},B}) - \ln \Omega_{\text{bath}}(E_{\text{bath},A})}{E_{\text{tot},B} - E_{\text{bath},A}} = \frac{1}{kT} \quad (7.1.5)$$

and this readily gives

$$\Omega_{\text{bath}}(E_{\text{tot}} - E_m) = \Omega_{\text{bath}}(E_{\text{tot}}) \exp\left(\frac{-E_m}{kT}\right) \quad (7.1.6)$$

and as E_{tot} is fixed, we have

$$P_m = B \exp\left(\frac{-E_m}{kT}\right) \quad (7.1.7)$$

where B is $1/q$, via the same normalisation argument.

7.1.2 Derivation of Boltzmann distribution

7.1.2.1 Via a Lagrange multiplier

We now derive the Boltzmann distribution law again, this time minimising a free energy. The problem we solve is essentially the same as the one in Section 6.1: we are given energy states E_j from the physics of the problem, and we aim to compute the probabilities p_j that the system is in j .

We now suppose that (T, V, N) are held constant, with only one species of particles. The

condition for equilibrium is then

$$dA = dU - T dS = 0. \quad (7.1.8)$$

This looks arbitrary but we can use any free energies as we have demonstrated in Section 6.5, minimising any of them under appropriate conditions is equivalent to maximising entropy.

Evidently now we need expressions for dS and dU . We have the expression for entropy:

$$S(\mathbf{p}) = -k \sum_{j=1}^t p_j \ln p_j, \quad (7.1.9)$$

which readily gives

$$dS = \sum_{p_i \neq j} \left(\frac{\partial S(\mathbf{p})}{\partial p_j} \right) = -k \sum_{j=1}^t (1 + \ln p_j) dp_j. \quad (7.1.10)$$

Postulate 7.1.2.1 (Internal energy). The internal energy is given as the average over all microscopic states:

$$U = \langle E \rangle = \sum_{j=1}^t p_j E_j. \quad (7.1.11)$$

We then have

$$dU = d\langle E \rangle = \sum_{j=1}^t (E_j dp_j + p_j dE_j). \quad (7.1.12)$$

Postulate 7.1.2.2 (Energy levels). Energy levels E_j only depends on volume and number of particles, and unlike U do not depend on S and T :

$$E_j = E_j(V, N). \quad (7.1.13)$$

This means

$$dE_j = \frac{\partial E_j}{\partial V} dV + \frac{\partial E_j}{\partial N} dN = 0. \quad (7.1.14)$$

As both V and N are held constant,

$$dE_j = 0, \quad (7.1.15)$$

and Equation 7.1.12 becomes

$$dU = d\langle E \rangle = \sum_{j=1}^t E_j dp_j. \quad (7.1.16)$$

Now we have our **objective function**:

$$dA = d\langle E \rangle - T dS = 0, \quad (7.1.17)$$

and the **constraint function** is then the usual

$$\begin{aligned} \sum_{j=1}^t p_j &= 1 \\ \sum_{j=1}^t dp_j &= 0. \end{aligned} \quad (7.1.18)$$

We can then write our Lagrange multiplier as

$$dA = \sum_{j=1}^t \left[E_j + kT(1 + \ln p_j^*) + \lambda \right] dp_j^* = 0. \quad (7.1.19)$$

This requires that

$$\ln p_j^* = -\frac{E_j}{kT} - \frac{\lambda}{kT} - 1 \quad (7.1.20a)$$

$$p_j^* = e^{-E_j/kT} e^{-(\lambda/kT)-1}. \quad (7.1.20b)$$

As p_j 's sum to 1,

$$\sum_{j=1}^t e^{-E_j/kT} e^{-(\lambda/kT)-1} = 1, \quad (7.1.21)$$

we can divide Equation 7.1.21 through Equation 7.1.20b to eliminate λ , and yield the

Theorem 7.1.2.1 (Boltzmann distribution law). The law states that

$$p_j^* = \frac{\exp(-E_j/kT)}{\sum_{j=1}^t \exp(-E_j/kT)} \equiv \frac{1}{q} \exp(-E_j/kT). \quad (7.1.22)$$

where q is the partition function and E_j is the energy of *state* j . When multiple states have the same energy *i.e.*, degenerate, it is convenient to merge them into *levels* of degeneracy g_i :

$$p_i = \frac{g_i}{q} \exp\left(\frac{-E_i}{kT}\right) \quad (7.1.23)$$

The Boltzmann distribution law is just a consequence of maximising entropy: given total energy, more molecules would have lower energy rather than a few having high energies and the rest having zero, to maximise total multiplicity.

7.1.2.2 From the canonical distribution

The Boltzmann distribution essentially arise from the canonical distribution when the particles are non-interacting, basically, we now can say that $Q_N = q^N/N!$ for indistinguishable particles, which is what we assume all along anyway. An outline of the proof is given below:

If we assume the particles are independent and non-interacting, we can write

$$Q_N = \frac{q^N}{N!} \quad (7.1.24)$$

and with

$$U = kT^2 \frac{\partial \ln Q_N}{\partial T} \quad (7.1.25)$$

we can write

$$U = NkT^2 \frac{1}{q} \left(\frac{\partial q}{\partial T} \right)_V \quad (7.1.26)$$

where

$$q = \sum_i \exp \left(\frac{-\epsilon_i}{kT} \right) \quad (7.1.27)$$

So

$$\left(\frac{\partial q}{\partial T} \right)_V = \frac{1}{kT^2} \sum_i \epsilon_i \exp \left(\frac{-\epsilon_i}{kT} \right) \quad (7.1.28)$$

Substituting this back into the expression for internal energy, we have

$$U = \frac{N}{q} \sum_i \exp \epsilon_i \exp \left(\frac{-\epsilon_i}{kT} \right) \quad (7.1.29)$$

But internal energy is also just the expectation value:

$$U = \sum_i n_i \epsilon_i \quad (7.1.30)$$

So clearly we can identify

$$\begin{aligned} n_i &= \frac{N}{q} \exp \left(\frac{-\epsilon_i}{kT} \right) \\ p_i &= \frac{1}{q} \exp \left(\frac{-\epsilon_i}{kT} \right) \end{aligned} \quad (7.1.31)$$

7.1.3 Interpretation of partition function

7.1.3.1 Physical meaning of the partition function

We introduced the partition function, but have not discussed why it's called 'partition function'. The partition function is

$$Q \equiv \sum_{j=1}^t e^{-E_j/kT}. \quad (7.1.32)$$

However, it's more commonly expressed in terms of energy differences as experiments often yield those. It is then convenient to define the ground-state as zero, $E_1 = 0$, and write

$$Q = 1 + e^{-(E_2-E_1)/kT} + e^{-(E_3-E_1)/kT} + \dots + e^{-(E_t-E_1)/kT}. \quad (7.1.33)$$

The two forms are equivalent.

Now, when the energies for all energy levels are small, *or* the temperature is high, all the states are equally populated and we have that

$$\frac{E_j}{kT} \rightarrow 0 \Rightarrow Q \rightarrow t \Rightarrow p_j^* \rightarrow \frac{1}{t}. \quad (7.1.34)$$

If all the *non-ground state* energies approach infinity, or the temperature approaches zero, the particles can only occupy the ground state:

$$\frac{E_{j \neq 1}}{kT} \rightarrow \infty \Rightarrow Q \rightarrow 1 \Rightarrow p_1^* \rightarrow 1. \quad (7.1.35)$$

In this case only the ground state becomes *effectively accessible*.

So we can see that the partition function gives the number of states that are *effectively* accessible to the system, and the magnitude of E_j/kT determines whether or not the state j is *effectively* accessible. But beware that the number of accessible states is always the same t and is independent of system variables as it is fixed by the physics of the situation. In contrast, Q is a function of system variables such as T .

7.1.3.2 Distinguishability

Definition 7.1.3.1 (Distinguishability). The distinction between distinguishable and indistinguishable particles lies in the de Broglie wavelength in comparison of typical particle separation of the type of particle in question, and *not in whether the particles are identical or not*.

For example, two tennis balls can be identical but occupy distinct locations and can thus be distinguished. The same is true for particles fixed in a lattice, which occupy a relatively fixed position.

However, two electrons in the He atom have de Broglie wavelengths about the same or larger than the interelectronic separation, hence cannot be distinguished. The same goes for particles in a gas, for example.

Now we consider systems made of two independent, distinguishable particles A and B . The energy levels are E_j and is given as a sum of individual energy levels:

$$E_j = \epsilon_i^A + \epsilon_m^B. \quad (7.1.36)$$

As the two particles are independent, they have independent partition functions,

$$q_A = \sum_{i=1}^a e^{-\epsilon_i^A/kT}, \quad (7.1.37)$$

and likewise for B .

The partition function for the system is the sum of Boltzmann factors $e^{-E_j/kT}$ over all $j = ab$ energy levels:

$$Q = \sum_{j=1}^t e^{-E_j/kT} = \sum_{i=1}^a \sum_{m=1}^b e^{-(\epsilon_i^A + \epsilon_m^B)/kT} = \sum_{i=1}^a \sum_{m=1}^b e^{-\epsilon_i^A/kT} e^{-\epsilon_m^B/kT}. \quad (7.1.38)$$

Because the two particles are distinguishable, all $j = ab$ energy levels are distinct and independent, for example, A in level 26 and B in level 53 is different from B in 26 and A in 53, hence we can write that

$$Q = \sum_{i=1}^a e^{-\epsilon_i^A/kT} \sum_{m=1}^b e^{-\epsilon_m^B/kT} = q_A q_B, \quad (7.1.39)$$

and in general, for a system with N distinguishable particles,

$$Q = q^N. \quad (7.1.40)$$

However, for indistinguishable particles, having A in level 23 and B in level 54 is exactly the same as the other way around, and the better way to specify this is to say that 'one

particle is in level 23 and another in level 54.¹ Therefore, if we were to count distinct energy levels, which is what the partition function is about, we should divide the count by $N!$ ¹:

$$Q = \frac{q^N}{N!}. \quad (7.1.41)$$

7.1.3.3 Gibbs's paradox

Another, and in my opinion, much better justification for dividing by $N!$ comes from [17] which shows that the division is merely a redefinition and has no fundamental justification, it is there to reconcile the statistical mechanical and thermodynamical definitions of entropy, the differences of which was revealed by the Gibbs's paradox, which provides for two isolated boxes with N identical particles and X accessible energy states, separated by an infinitely thin and isolating division. Let's call the system before and after removing the division A and B , and the multiplicities are given as

$$S_A = X^{2N} \quad (7.1.42a)$$

$$S_B = (2X)^{2N}. \quad (7.1.42b)$$

The change in entropy is

$$S_B - S_A = 2kN \ln 2, \quad (7.1.43)$$

but the change must be

$$S_B - S_A = \int_A^B \frac{dQ}{T}, \quad (7.1.44)$$

as the gases remain in equilibrium the removal of division is a quasistatic process so the change *must* be zero. This is the Gibbs's paradox. To resolve it, we must realise that the definition of

$$S = k \ln W \quad (7.1.45)$$

is for statistical mechanics and not for thermodynamics, and that, indeed the number of accessible energy levels has increased but this does not correspond well with the usual notion of entropy, hence we need to redefine entropy as

$$S = k \ln \frac{W}{N!}. \quad (7.1.46)$$

¹This is an approximation as if the particles coincide on the same energy level, there will not be over-counting, but this is in reality often negligible as the number of accessible levels are usually much larger than number of particles.

Now, the ‘new’ multiplicities are

$$W_A = \frac{X^{2N}}{(N!)^2} \quad (7.1.47a)$$

$$W_B = \frac{(2X)^{2N}}{(2N)!}. \quad (7.1.47b)$$

We apply Stirling’s approximation under the assumption of the *thermodynamic limit* of $N \rightarrow \infty$:

$$\ln n! \approx n \ln n - n \quad (7.1.48)$$

to show that

$$\ln W_A = 2N \ln X - 2 \ln N! \approx 2N \ln X - 2N \ln N + 2N \quad (7.1.49a)$$

$$\ln W_B = 2N \ln 2X - 2 \ln(2N)! \approx 2N \ln 2X - 2N \ln 2N + 2N. \quad (7.1.49b)$$

The two expressions are equal, and we have resolved the Gibb’s paradox. This also provides another justification for dividing the partition function by $N!$.

7.1.4 Applications of the Boltzmann distribution

7.1.4.1 Population of states

We go through partition functions in great detail in Section 7.2, and we use the results from there:

Populations of vibrational energy levels

For a simple harmonic oscillator, the energy levels, with the lowest level set to 0, are

$$\epsilon_\nu = \nu h\nu_0 \quad (7.1.50)$$

and

$$q'_{\text{vib}} = \frac{1}{1 - \exp(-h\nu_0/kT)} \quad (7.1.51)$$

so we have

$$p_\nu = \left[1 - \exp\left(\frac{-\theta_{\text{vib}}}{T}\right) \right] \exp\left(\frac{-\nu\theta_{\text{vib}}}{T}\right) \quad (7.1.52)$$

Populations of rotational energy levels

For a rigid rotor, the levels are

$$\epsilon_J = BJ(J+1) \quad (7.1.53)$$

and the levels have degeneracies of $(2J+1)$, so

$$p_J = \frac{2J+1}{q_{\text{rot}}} \exp\left(\frac{-BJ(J+1)}{kT}\right) \quad (7.1.54)$$

In the high temperature limit (reached at very low actual temperatures, so almost always valid),

$$q_{\text{rot}} = \frac{T}{\sigma\theta_{\text{rot}}} \quad (7.1.55)$$

where $\theta_{\text{rot}} = B/k$, we have

$$p_J = \frac{\sigma\theta_{\text{rot}}}{T} (2J+1) \exp\left(\frac{-\theta_{\text{rot}}J(J+1)}{T}\right) \quad (7.1.56)$$

The degeneracy term wins out the exponential form at small J and so intensity increases at first, but eventually tails out. The maximum can be found by setting $dp_J/dJ = 0$, which gives

$$J_{\text{max}} = \sqrt{\frac{T}{2\theta_{\text{rot}}}} - \frac{1}{2} \quad (7.1.57)$$

7.1.4.2 Density of states

The density of states $D(\epsilon)$ is the number of states per energy, or $D(\epsilon) d\epsilon$ is the number of states between ϵ to $\epsilon + d\epsilon$. It is formally the derivative of number of states, $W(\epsilon)$, which gives the number of states up to ϵ :

$$D(\epsilon) = \frac{dW(\epsilon)}{d\epsilon} \quad (7.1.58)$$

We consider the very densely packed case of translational energy states: in a cubical box of side a , the energy levels are

$$E_{n_x, n_y, n_z} = (n_x^2 + n_y^2 + n_z^2) \frac{h^2}{8ma^2} \quad (7.1.59)$$

This is reminiscent of the Cartesian expression for a sphere, with $n_x^2 + n_y^2 + n_z^2$ the squared radius, so the 'radius' is given as

$$n_{\max} = \left(\frac{8ma^2\epsilon}{h^2} \right)^{1/2} \quad (7.1.60)$$

However, only the octant corresponding to positive (x, y, z) can be taken as n_i 's are nonzero, so the number of states are

$$W(\epsilon) = \frac{1}{8} \times \frac{4}{3} \pi n_{\max}^3 = \frac{\pi}{6} \left(\frac{8m\epsilon}{h^2} \right)^{3/2} V \quad (7.1.61)$$

where we wrote a^3 as V . $D(\epsilon)$ soon follows:

Theorem 7.1.4.1 (Density of translational states in a box). For a cubical box with volume V , $D(\epsilon)$ is given as

$$D(\epsilon) = \frac{dW(\epsilon)}{d\epsilon} = \frac{\pi}{4} \left(\frac{8m}{h^2} \right)^{3/2} \epsilon^{1/2} V \quad (7.1.62)$$

This is a result that will come in handy when we consider the free electron gas model of metal bonding.

7.1.4.3 Non-equilibrium states

The population ratio between states is given as

$$\frac{n_1}{n_0} = \exp \left(\frac{-\Delta\epsilon}{kT} \right) \quad (7.1.63)$$

It is possible to excite population from the lower energy level selectively by methods such as intense radiation, such that $n_1 > n_0$ and the temperature is temporarily negative. This is called a *population inversion*, which is important in the working of lasers.

7.1.4.4 Maxwell-Boltzmann distribution of velocities

We adopt the assumptions of the kinetic theory that gas particles can be treated as perfect Newtonian particles. The kinetic energy, which is the total energy is given by

$$\epsilon(v) = \frac{1}{2}mv^2. \quad (7.1.64)$$

According to the Boltzmann law, the probability $p(v_x)$ that a particle in a container at constant volume and temperature will have velocity v_x in the x direction is

$$p(v_x) = \frac{e^{-\epsilon/kT}}{\int_{-\infty}^{+\infty} e^{-\epsilon/kT} dv_x}, \quad (7.1.65)$$

where dv_x serves the same role as the rolling index j in Equation 7.1.22. We write further that

$$p(v_x) = \frac{e^{-mv_x^2/2kT}}{\int_{-\infty}^{+\infty} e^{-mv_x^2/2kT} dv_x} \quad (7.1.66a)$$

$$= \left(\frac{m}{2\pi kT} \right)^{1/2} e^{-mv_x^2/2kT} \quad (7.1.66b)$$

The *mean-square* velocity is

$$\langle v_x^2 \rangle = \int_{-\infty}^{+\infty} v_x^2 p(v_x) dv_x \quad (7.1.67a)$$

$$= \left(\frac{m}{2\pi kT} \right)^{1/2} \int_{-\infty}^{+\infty} v_x^2 e^{-mv_x^2/2kT} dv_x \quad (7.1.67b)$$

$$= \left(\frac{m}{2\pi kT} \right)^{1/2} \left(\frac{kT}{m} \right) \left(\frac{2\pi kT}{m} \right)^{1/2} \quad (7.1.67c)$$

$$= \frac{kT}{m}. \quad (7.1.67d)$$

This means that the average kinetic energy is

$$\frac{1}{2} m \langle v_x^2 \rangle = \frac{1}{2} kT. \quad (7.1.68)$$

In three dimensions, because

$$|\mathbf{v}|^2 = \mathbf{v} \cdot \mathbf{v} = v_x^2 + v_y^2 + v_z^2, \quad (7.1.69)$$

and because all three components are treated as independent,

$$\langle v^2 \rangle = \langle v_x^2 \rangle + \langle v_y^2 \rangle + \langle v_z^2 \rangle, \quad (7.1.70)$$

so

$$\frac{1}{2} \langle v^2 \rangle = \frac{3}{2} kT. \quad (7.1.71)$$

As the velocities are independent, the probabilities multiply to give

$$p(v) = p(v_x)p(v_y)p(v_z) = \left(\frac{m}{2\pi kT} \right)^{3/2} e^{-mv^2/2kT}. \quad (7.1.72)$$

7.1.5 Thermodynamic properties from partition functions

7.1.5.1 Distinguishable particles

Internal energy

For a system with fixed (T, V, N) , to get the internal energy, we remember

$$U = \langle E \rangle = \sum_{j=1}^t p_j E_j \quad (7.1.73a)$$

$$= \frac{1}{Q} \sum_{j=1}^t E_j e^{-\beta E_j}, \quad (7.1.73b)$$

where $\beta = 1/kT$. We note that the sum in the last equality is a derivative of the partition function:

$$\frac{\partial Q}{\partial \beta} = \frac{\partial}{\partial \beta} \sum_{j=1}^t e^{-\beta E_j} = - \sum_{j=1}^t E_j e^{-\beta E_j}. \quad (7.1.74)$$

Therefore,

$$U = - \frac{1}{Q} \frac{\partial Q}{\partial \beta} = - \frac{\partial \ln Q}{\partial \beta} = - \frac{\partial \ln Q}{\partial T} \frac{\partial T}{\partial \beta} = kT^2 \frac{\partial \ln Q}{\partial T} \quad (7.1.75)$$

Average particle energy

For distinguishable particles we have $Q = q^N$ so

$$\langle \epsilon \rangle = \frac{U}{N} = \frac{kT^2}{N} \frac{\partial \ln q^N}{\partial T} = kT^2 \frac{\partial \ln q}{\partial T}. \quad (7.1.76)$$

Entropy

As entropy is defined as

$$S = -k \sum_{j=1}^t p_j \ln p_j, \quad (7.1.77)$$

we can get

$$S = -k \sum_{j=1}^t \left(\frac{1}{Q} e^{-E_j/kT} \right) \left[\ln \left(\frac{1}{Q} \right) - \frac{E_j}{kT} \right] \quad (7.1.78a)$$

$$= -k \left[\sum_{j=1}^t \left(\frac{1}{Q} e^{-E_j/kT} \right) \ln \left(\frac{1}{Q} \right) - \sum_{j=1}^t \left(\frac{1}{Q} e^{-E_j/kT} \right) \frac{E_j}{kT} \right] \quad (7.1.78b)$$

$$= -k \frac{Q}{Q} \ln \left(\frac{1}{Q} \right) + \frac{U}{T} \quad (7.1.78c)$$

$$= k \ln Q + \frac{U}{T} \quad (7.1.78d)$$

Free energies and potentials

We have

$$A \equiv U - TS. \quad (7.1.79)$$

So obviously,

$$A = -kT \ln Q. \quad (7.1.80)$$

And more definitions follow:

$$\mu = \left(\frac{\partial A}{\partial N} \right)_{T,V} = -kT \left(\frac{\partial \ln Q}{\partial N} \right)_{T,V}, \quad (7.1.81)$$

and so on.

7.1.5.2 Indistinguishable particles

In this case $Q_N = q^N/N!$, and under Stirling's approximation

$$\ln N! = N \ln N - N, \quad (7.1.82)$$

we can write...

Internal energy

$$\begin{aligned} U &= kT^2 \left(\frac{\partial \ln q^N/N!}{\partial T} \right)_V \\ &= NkT^2 \left(\frac{\partial \ln q}{\partial T} \right)_V \end{aligned} \quad (7.1.83)$$

Entropy

$$\begin{aligned} S &= k \ln \frac{q^N}{N!} + \frac{U}{T} \\ &= Nk \ln q - Nk \ln N + Nk + \frac{U}{T} \end{aligned} \quad (7.1.84)$$

Free energy

$$\begin{aligned}
 A &= -kT \ln \frac{q^N}{N!} \\
 &= -NkT \ln q + kT(N \ln N - N)
 \end{aligned}
 \tag{7.1.85}$$

Chemical potential

$$\mu = \left(\frac{\partial A}{\partial N} \right)_{T,V} = -kT \ln \left(\frac{q}{N} \right) \tag{7.1.86}$$

7.1.6 Examples of the two-state model

7.1.6.1 The Schottky model

Consider a system with N distinguishable particles with just two energy levels for each particle: a ground state with energy zero and an excited state with energy $\epsilon = \epsilon_0 > 0$. This is a widely applicable model such as our polymer or dimer models, or in general any excitable particle systems.

Here we give a general discussing, aiming to find the average energy $\langle \epsilon \rangle$, the heat capacity C_V , the entropy and the free energy per particle, from the partition function, which is given by

$$q = 1 + e^{-\beta \epsilon_0}. \tag{7.1.87}$$

q approaches 1 at low temperatures and 2 at high temperatures, with probabilities given by

$$p_1^* = \frac{1}{q} \quad \text{and} \quad p_2^* = \frac{e^{-\beta \epsilon_0}}{q}. \tag{7.1.88}$$

Average energy is

$$\langle \epsilon \rangle = \sum p_j^* \epsilon_j = \epsilon_0 p_2^* = \frac{\epsilon_0 e^{-\epsilon_0/kT}}{1 + e^{-\epsilon_0/kT}}. \tag{7.1.89}$$

Heat capacity is given by

$$C_V = \left(\frac{\partial U}{\partial T} \right)_{V,N} = N \left(\frac{\partial \langle \epsilon \rangle}{\partial T} \right)_{V,N}. \tag{7.1.90}$$

The evaluation of the derivative above is not straightforward, we can make it simpler by writing

$$N \left(\frac{\partial \langle \epsilon \rangle}{\partial T} \right)_{V,N} = N \frac{\partial \langle \epsilon \rangle}{\partial \beta} \frac{d\beta}{dT} = \frac{N \epsilon_0^2}{kT^2} \frac{e^{-\beta \epsilon_0}}{(1 + e^{-\beta \epsilon_0})^2}. \tag{7.1.91}$$

As shown in Figure 7.1 (b), there is a characteristic sharp peak in C_V at intermediate temperatures, when the energy of the heat bath can most effectively excite the particles in

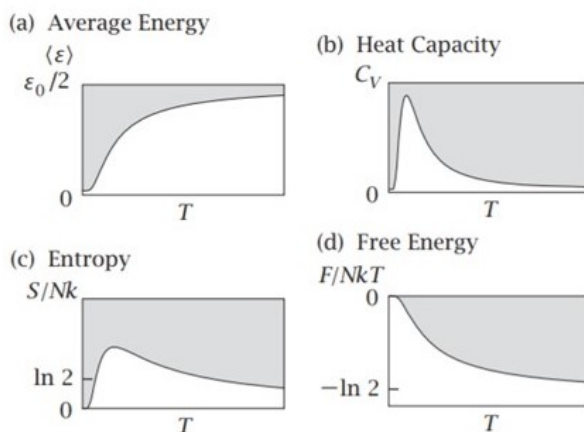


Figure 7.1

the system.

Entropy is obtained by remembering

$$S = k \ln Q + \frac{U}{T} \quad (7.1.92a)$$

$$= Nk \ln(1 + e^{-\epsilon_0/kT}) + \frac{N\epsilon_0 e^{-\epsilon_0/kT}}{T(1 + e^{-\epsilon_0/kT})}, \quad (7.1.92b)$$

and lastly, free energy:

$$A = -kT \ln Q = -NkT \ln(1 + e^{-\epsilon_0/kT}). \quad (7.1.93)$$

As $\epsilon_0 \rightarrow \infty$, the excited states becomes inaccessible and so $S \rightarrow 0$ and $A \rightarrow 0$.

As $\epsilon_0 \rightarrow 0$, two states are equally accessible and so $S \rightarrow Nk \ln 2$ and $A \rightarrow -NkT \ln 2$.

7.1.6.2 Curie's law of paramagnetism

Paramagnetic materials are made up of atoms which have negligible dipole-dipole as compared to interaction with an applied field. Each atom has a magnetic dipole of magnitude $\mu_0 > 0$, and gives an energy of $-\mu_0 B$ if the spin is aligned parallel to the field and $+\mu_0 B$ if aligned antiparallel. Setting the ground state energy to 0 we have our partition function:

$$q = 1 + e^{-2\mu_0 B/kT}. \quad (7.1.94)$$

We would like to calculate the average magnetic moment. The probabilities are

$$p_1^* = \frac{1}{q} \quad \text{and} \quad p_2^* = \left(\frac{1}{q}\right) e^{-2\mu_0 B/kT}. \quad (7.1.95)$$

So the average dipole moment is calculated by

$$\langle \mu \rangle = \mu_0 p_1^* + (-\mu_0) p_2^* = \mu_0 \frac{1 - e^{-2\mu_0 B/kT}}{1 + e^{-2\mu_0 B/kT}} = \mu_0 \tanh \left(\frac{\mu_0 B}{kT} \right). \quad (7.1.96)$$

The Taylor series expansion gives

$$\langle \mu \rangle = \frac{\mu_0^2 B}{kT} \quad (7.1.97)$$

at small $\mu_0 B/kT$, meaning high temperature and/or weak field. This is **Curie's law**. This explains the fact that cooling and high field strengths magnetise paramagnets effectively.

7.1.6.3 Modelling a protein loop

test

7.2 THE STATISTICAL MECHANICS OF SIMPLE GASES

7.2.1 Molecular partition function

7.2.1.1 Translational partition function

The quantum mechanical model of a particle in a potential well provides energy levels for such a particle, and the energies are given by

$$\epsilon_n = \frac{nh^2}{8mL^2}, \quad (7.2.1)$$

where L is the length of the well. So the translational partition function is given by

$$q_{\text{trans}} = \sum_{n=1}^{\infty} e^{-\epsilon_n/kT} = \sum_{n=1}^{\infty} e^{-n^2 h^2 / (8mL^2 kT)}. \quad (7.2.2)$$

It is customary to define

Definition 7.2.1.1 (Temperature of degrees of freedom). Energies associated with degrees of freedom are usually divided by the Boltzmann factor and expressed in units of temperature. For example, the **translational temperature** is

$$\theta_{\text{trans}} = \frac{h^2}{8mL^2 k}, \quad (7.2.3)$$

such that partition function can be compactly written as

$$q_{\text{trans}} = \sum_{n=1}^{\infty} e^{-n^2 \theta/T}, \quad (7.2.4)$$

If the temperature of a degree of freedom is low, then the partition function can be *approximated by an integral*.

Evidently, when $\theta/T \ll 1$, the ability of the n^2 factor of bringing the magnitude of the exponential is limited, and a great many Boltzmann's factors will be evaluated at around 1 until very large n is reached. This means the partition function, and hence the number of effectively populated states is very large. This means we can approximate the sum as

an integral:

$$q_{\text{trans}} = \int_0^\infty e^{(-h^2/8mL^2kT)n^2} dn = \left(\frac{2\pi mkT}{h^2} \right)^{1/2} L \equiv \frac{L}{\Lambda}, \quad (7.2.5)$$

where $\Lambda = (2\pi mkT/h^2)^{-1/2}$ is the **thermal wavelength**. Now, for a 3D box of side lengths a, b, c , as the motion in each direction is assumed to be *independent*, the wavefunction $\psi(x, y, z)$ can be separated into $\psi(x)\psi(y)\psi(z)$, and we can write

$$H\psi(x, y, z) = (H_x + H_y + H_z)\psi(x)\psi(y)\psi(z) = \epsilon_{n_x, n_y, n_z}\psi(x, y, z), \quad (7.2.6)$$

where

$$H \equiv -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + \underbrace{V(x, y, z)}_{=0}, \quad (7.2.7)$$

and

$$\epsilon_{n_x, n_y, n_z} = \frac{h^2}{8m} \left(\frac{n_x^2}{a^2} + \frac{n_y^2}{b^2} + \frac{n_z^2}{c^2} \right). \quad (7.2.8)$$

As the energy can be separated into three dimensions (a direct result of the fact that their motions are independent, anyway), the Boltzmann factors are multiplied and hence can be independently evaluated and so the resulting partition functions are just the product of three one-dimensional partition functions:

$$q_{\text{trans}} = q_x q_y q_z = \left(\frac{2\pi mkT}{h^2} \right)^{3/2} V \equiv \frac{V}{\Lambda^3} \quad (7.2.9)$$

Worked example 7.2.1.1. For temperature $T = 273\text{K}$, and the volume $V = 2.24 \times 10^{-2}\text{m}^3$ of one mole of gas at $p = 1\text{atm}$, the translational partition function is

$$q_{\text{trans}} = \left(\frac{2\pi mkT}{h^2} \right)^{3/2} V. \quad (7.2.10)$$

Substituting in mass for a single atom computed from molar mass, we have

$$q_{\text{trans}} \approx 4.79 \times 10^{30} \text{ states per atom.} \quad (7.2.11)$$

7.2.1.2 Vibrational partition function

The quantum mechanical harmonic oscillator model gives the energy of a harmonic oscillator as

$$\epsilon_v = \hbar\omega \left(v + \frac{1}{2} \right), \quad v = 0, 1, 2, \dots, \quad (7.2.12)$$

where $\omega = \sqrt{k/\mu}$. We set the ground state energy $\epsilon_0 = \hbar\omega/2$ to 0, to get the partition function

$$q_{\text{vib}} = \sum_{v=0}^{\infty} e^{-\hbar\omega v/kT} \equiv 1 + x + x^2 + \cdots = (1-x)^{-1} = \frac{1}{1 - e^{-\hbar\omega/kT}}. \quad (7.2.13)$$

We can use the Taylor series expansion because the vibrational temperature, given by

$$\theta_{\text{vib}} = \frac{\hbar\omega}{k} \approx 10^3 \text{ K} \quad (7.2.14)$$

is large enough to make sure that $|x| < 1$, which is the condition of convergence for this Taylor series.

Worked example 7.2.1.2. Oxygen molecules have a vibrational wavenumber of 1580 cm^{-1} , so

$$\theta_{\text{vib}} = \frac{\hbar\omega}{k} \approx 2274 \text{ K}. \quad (7.2.15)$$

So the partition function at $T = 300 \text{ K}$ is

$$q_{\text{vib}} = \frac{1}{1 - e^{-\theta_{\text{vib}}/T}} \approx 1.0005. \quad (7.2.16)$$

So most oxygen molecules are in their ground vibrational states at this temperature. And even at 1000 K , the partition function is only 1.11.

For more complex molecules, as each vibrational mode has a separate set of vibrational wavefunctions and energies associated with them, the total partition function is just a product of the partition functions of each of the modes.

7.2.1.3 Rotational partition function

The rotational energy is given by

$$\epsilon_{\ell} = \frac{\hbar^2}{2I} \ell(\ell+1), \quad \ell = 0, 1, 2, \dots \quad (7.2.17)$$

There is a $2\ell + 1$ -fold degeneracy thanks to m , so the partition function is

$$q_{\text{rot}} = \sum_{\ell=0}^{\infty} (2\ell+1) e^{-\epsilon_{\ell}/kT} \equiv \sum_{\ell=0}^{\infty} (2\ell+1) e^{-\theta_{\text{rot}}/T}. \quad (7.2.18)$$

The rotational temperature is defined as

$$\theta_{\text{rot}} \equiv \frac{\hbar^2}{2Ik} = \frac{\tilde{B}}{k}, \quad (7.2.19)$$

and is typically in the order of 10 K, therefore, only in *high temperature limit*, can we approximate the sum as an integral:

$$q_{\text{rot}} = \int_0^{\infty} (2\ell + 1) e^{-(\ell^2 + \ell)\theta/T} d\ell = \frac{T}{\theta}. \quad (7.2.20)$$

All very well, but we have to introduce a rotational symmetry factor σ :

$$q_{\text{rot}} \equiv \frac{T}{\sigma\theta}. \quad (7.2.21)$$

This factor accounts for the ways the molecule can indistinguishably overlap itself by rotation, for example, it takes on the values

$$\sigma = \begin{cases} 1, & \text{heteronuclear diatomics} \\ 2, & \text{homonuclear diatomics} \\ 2, & \text{H}_2\text{O} \\ 12, & \text{CH}_4 \\ 12, & \text{benzene.} \end{cases} \quad (7.2.22)$$

These are essentially the number of all rotational classes in the point group the molecule belongs to, plus one (E operation). For nonlinear molecules with three principal moments of inertia I_a , I_b and I_c , the partition function is given, in the high temperature limit (**proof needed**):

$$q_{\text{rot}} = \frac{(\pi I_a I_b I_c)^{1/2}}{\sigma} \left(\frac{2kT}{\hbar^2} \right)^{3/2} = \frac{\pi^{1/2}}{\sigma} \left(\frac{T^3}{\theta_a \theta_b \theta_c} \right)^{1/2}, \quad (7.2.23)$$

where the a, b, c subscripts stand for the three axes of rotation. However, for smaller molecules or lower temperatures, we cannot justifiably use the integral and must use the **Euler-Maclaurin formula** to approximate the sum.

$$q_{\text{rot}} = \frac{T}{\theta} + \frac{1}{3} + \frac{1}{15} \frac{\theta}{T} + \frac{4}{315} \left(\frac{\theta}{T} \right)^2 + \frac{1}{315} \left(\frac{\theta}{T} \right)^3 + \dots \quad (7.2.24)$$

how to get this from E-M formula?

7.2.1.4 Electronic partition function

The electronic energy levels are usually large compared to kT and as a result the excited levels do not contribute significantly to the overall partition function, however exceptions do exist, for example for Si(g)

| level | 3P_0 | 3P_1 | 3P_2 | 1D_2 | 1S_0 |
|---------------------------|---------|---------|---------|---------|---------|
| energy / cm^{-1} | 0.0 | 76 | 222 | 6297 | 15390 |

The electronic partition function q_{el} is given by

$$q_{\text{el}} = \sum_i g_i \exp(-\epsilon_i/kT), \quad (7.2.25)$$

where $g_i = 2J_i + 1$, as the *levels* (J) are degenerate (can take on different M_J , which are called states) in absense of external magnetic fields. In this case q_{el} is about 6.05.

In **molecules**, the degeneracy is the product of the spin and spatial degeneracies, for example, $^3\Pi_u$ is $3 \times 2 = 6$ degenerate.

7.2.2 Interal energy and heat capacity

An rule of thumb that works when $T \gg \theta$ is the equipartion principle.

Theorem 7.2.2.1 (Equipartition principle). The equipartition principle states that each 'squared term' in the expression for the total energy^a contributes $\frac{1}{2}RT$ per mole to the internal energy of the system.

^aQuadratic in position or momentum variables

To break it down, we look at the four contributors to internal energy.

Translations

Three squared terms, contribute to $3RT/2$ per mole.

Rotations

Three squared terms for *non-linear* molecules, but two for *linear* molecules. $3RT/2$ or RT contribution respectively.

Vibrations

Vibrational energy is a sum of the kinetic energy and potential energy, therefore two squared terms are involved and contributes RT to the internal energy.

Electronic states

It does not contain a squared term and falls outside the remit of the equipartition principle, but the ground state electronic energy is a significant contributor to internal energy **update this**

As $C_v = (\partial U / \partial T)_V$, we can easily see that the heat capacity will vary as follows

| T / K | linear | non-linear |
|---|----------------|----------------|
| $T < \theta_{\text{rot}}$ | $\frac{3}{2}R$ | $\frac{3}{2}R$ |
| $\theta_{\text{rot}} < T < \theta_{\text{vib}}$ | $\frac{5}{2}R$ | $3R$ |
| $T > \theta_{\text{vib}}$ | $\frac{7}{2}R$ | $4R$ |

However the equipartition principle only works at large temperatures for each degree of freedom ($T \gg \theta$), and we outline the proof below **Translational contribution**

Starting from

$$q_{\text{trans}} = \left(\frac{2\pi mkT}{h^2} \right)^{3/2} V \equiv \alpha T^{3/2}, \quad (7.2.26)$$

which *already assumes the high temperature limit*, and

$$U = NkT^2 \left(\frac{\partial \ln q}{\partial T} \right)_V \quad (7.2.27)$$

we arrive at

$$\begin{aligned} U_{\text{trans}} &= NkT^2 \left(\frac{\partial \ln \alpha T^{3/2}}{\partial T} \right)_V \\ &= NkT^2 \left(\frac{\partial \left(\ln \alpha + \frac{3}{2} \ln T \right)}{\partial T} \right)_V \\ &= \frac{3}{2} NkT, \end{aligned} \quad (7.2.28)$$

and also that $C_{V,m,\text{trans}} = 3R/2$.

Rotational contribution

Starting from

$$q_{\text{rot}} = \frac{T}{\sigma \theta_{\text{rot}}}, \quad (7.2.29)$$

which *also assumes the high temperature limit*, we get

$$\begin{aligned} U_{\text{rot}} &= NkT^2 \left(\frac{\partial \ln q_{\text{rot}}}{\partial T} \right)_V \\ &= NkT^2 \left(\frac{\partial \ln \alpha T}{\partial T} \right)_V \\ &= NkT \end{aligned} \quad (7.2.30)$$

Which means a contribution of R to the heat capacity.

Vibrational contribution

We start from

$$q'_{\text{vib}} = \frac{1}{1 - \exp(-\theta_{\text{vib}}/T)}, \quad (7.2.31)$$

which *does not assume* high temperature limit, which we'll have to do later to recover the equipartition principle:

$$\begin{aligned} U'_{\text{vib}} &= NkT^2 \left(\frac{\partial \ln q'_{\text{vib}}}{\partial T} \right)_V \\ &\dots \\ &= \frac{Nk\theta_{\text{vib}}}{\exp(\theta_{\text{vib}}/T) - 1} \end{aligned} \quad (7.2.32)$$

C_v is tedious to compute

$$C_{v,\text{vib}} = \frac{Nk\theta_{\text{vib}}^2}{T^2} \frac{\exp(\theta_{\text{vib}}/T)}{[\exp(\theta_{\text{vib}}/T) - 1]^2}. \quad (7.2.33)$$

In the high temperature limit the Taylor expansion $e^x \approx 1 + x + x^2/2 + \dots$ applies,

$$\begin{aligned} q'_{\text{vib}} &\approx \frac{T}{\theta_{\text{vib}}} \\ U'_{\text{vib}} &\approx NkT \\ C_v &\approx Nk \end{aligned} \quad (7.2.34)$$

7.2.3 Entropy

We can write the total partition function for a system with N particles as

$$\begin{aligned} Q &= Q_{\text{trans}} Q_{\text{rot}} Q_{\text{vib}} Q_{\text{el}} \\ &= \frac{q_{\text{trans}}}{N!} q_{\text{rot}}^N q_{\text{vib}}^N q_{\text{el}}^N, \end{aligned} \quad (7.2.35)$$

where the indistinguishability of the particles only needs to be taken into account *once*, and rotations etc. are treated as if the particles were in a crystal *i.e.*, distinguishable. This really only applies when the particles are *weakly coupled*, which is not the case in liquids and solids.

With that out of the way, we recall that

$$S = k \ln Q + \frac{U}{T}. \quad (7.2.36)$$

Translational entropy

As

$$q_{\text{trans}} = \left(\frac{2\pi mkT}{h^2} \right)^{3/2} V, \quad (7.2.37)$$

we get

Theorem 7.2.3.1 (The Sackur-Tetrode Equation). The equation states

$$\begin{aligned} S_{\text{trans}} &= Nk \ln \left[\left(\frac{2\pi mkT}{h^2} \right)^{3/2} V \right] - Nk \ln N + Nk + \frac{3}{2} Nk \\ &= Nk \left\{ \ln \left[\left(\frac{2\pi mkT}{h^2} \right)^{3/2} V \right] - \ln N + \frac{5}{2} \right\} \end{aligned} \quad (7.2.38)$$

Further simplification under molar ideal gas conditions ($N = N_A$, $Nk = R$, $V = RT/p$, $m = M/N_A$)

$$\begin{aligned} S_{\text{trans},m} &= R \left[\ln \left(\frac{M^{3/2} T^{5/2}}{p} \right) \right] + 20.723R \\ &= \underbrace{R \ln V + \frac{3}{2} R \ln T}_{\text{classical}} + \underbrace{\frac{3}{2} R \ln M + 18.605R}_{\text{non-classical}} \end{aligned} \quad (7.2.39)$$

Rotational entropy

$$q_{\text{rot}} = \frac{T}{\sigma \theta_{\text{rot}}}, \quad (7.2.40)$$

so

$$\begin{aligned} S_{\text{rot}} &= Nk \ln q_{\text{rot}} + \frac{U_{\text{rot}}}{T} \\ &= Nk \left[\ln \left(\frac{T}{\sigma \theta_{\text{rot}}} \right) + 1 \right] \end{aligned} \quad (7.2.41)$$

remembering the different values of U_{vib} for linear and non-linear molecules, so for a non-linear molecule

$$S_{\text{rot}} = Nk \left[\ln \left(\frac{T}{\sigma \theta_{\text{rot}}} \right) + \frac{3}{2} \right] \quad (7.2.42)$$

Vibrational entropy

Not much can be done to simplify this, just substitute in eqs. 7.2.31 and 7.2.32,

$$S_{\text{vib}} = Nk \ln q'_{\text{vib}} + \frac{U'_{\text{vib}}}{T} \quad (7.2.43)$$

Often it is the case that $q'_{\text{vib}} = 1$ and so $U'_{\text{vib}} = 0$, and there's no vibrational contribution, as to be expected when all particles are in ground state.

Electronic entropy

$$S_{\text{el}} = Nk \ln q'_{\text{el}} + \frac{U'_{\text{el}}}{T} = Nk \ln g_0 \quad (7.2.44)$$

Calorimetry

To determine entropy experimentally, the following expression is used

$$\begin{aligned} S(T) &= \int_0^T \frac{C_p(T')}{T'} dT' + \sum_{\text{pc}} \frac{\Delta H_{\text{pc}}}{T_{\text{pc}}} \\ &\approx \underbrace{\int_0^{T_0} \frac{aT'^3}{T'} dT'}_{\text{Debye law}} + \int_{T_0}^T \frac{C_p(T')}{T'} dT' + \sum_{\text{pc}} \frac{\Delta H_{\text{pc}}}{T_{\text{pc}}} \end{aligned} \quad (7.2.45)$$

7.2.4 Nuclear spin partition function

The nuclear spin quantum number I can be an integer or a half-integer, and the projection onto z axis is M_I , which run from $-I$ to I . In absence of external magnetic field these levels are degenerate.

In a diatomic molecule, in the weak-coupling approximation there are $(2I_1 + 1) \times (2I_2 + 1)$.

Now considering the example of $^1\text{H}^{19}\text{F}$, both have $I = 1/2$ so we can denote the possible states as

$$\alpha(H)\alpha(F), \alpha(H)\beta(F), \beta(H)\alpha(F), \beta(H)\beta(F), \quad (7.2.46)$$

For a homonuclear diatomic though as the two nuclei are identical the situation is more complex, as for example $\alpha(1)\beta(2)$ produces $\alpha(2)\beta(1)$ is produced which is a completely different function which is unacceptable. We can follow the example of electron spins and product three triplet states, that *doesn't* change sign under exchange of labels, and one singlet state, which *does* change sign. **incomplete**

Theorem 7.2.4.1 (Number of sym and antisym nuclear spin states). For a homonuclear diatomic, each atom with the nuclear spin I

$$\begin{aligned} \# \text{ of symmetric states} &= (2I + 1)(I + 1) \\ \# \text{ of antisymmetric states} &= (2I + 1)I \end{aligned} \quad (7.2.47)$$

Proof. There are $(2I + 1)^2$ states in total, of which, $(2I + 1)$ states have both the nuclei with the same spin *i.e.*, symmetric. The remaining states aren't simply symmetric or antisymmetric as nuclear exchange produces entirely different states. Therefore two linear

combinations can be formed between two permuted states² to produce symmetric and antisymmetric wavefunctions. This means half of the remaining states will be symmetric and half antisymmetric. This brings the total to

$$\begin{aligned}\text{\# of symmetric states} &= (2I + 1) + (2I + 1)I = (2I + 1)(I + 1) \\ \text{\# of antisymmetric states} &= (2I + 1)I\end{aligned}\tag{7.2.48}$$

□

7.2.4.1 Nuclear exchange

Exchange of nuclear labels must be done in way that does not affect other wavefunctions, so a simple 180° rotation (C_2) won't do. The following operations are needed

1. C_2 , but electrons around the nuclei are moved as well, nuclear spin states are not affected.
2. i_{elec} Invert just the electrons through the centre of inversion.
3. σ_{elec} Reflect just the electrons in a plane containing the internuclear axis.
4. ρ_{nuc} , permutes (flips) the nuclear spin states.

These are going to have the following effects on the overall wavefunction

1. C_2 : ψ_{elec} and ψ_{vib} are unaffected, however ψ_{rot} will be affected in the same way as the spherical harmonics, which means even J is unchanged, and odd J changes sign
2. i_{elec} : this depends on the inversion symmetry (g/u) of the molecular term symbol, g doesn't change sign and u does.
3. σ_{elec} : this depends on the reflection symmetry (\pm), $+$ doesn't change sign and $-$ does.
4. ρ_{nuc} : symmetric states don't change sign and asymmetric states do.

7.2.4.2 Consequences on spectroscopy

We will look at the consequence of this for a few examples.

²This means $\alpha(1)\gamma(2)$ and $\gamma(1)\alpha(2)$, or $\delta(1)\beta(2)$ and $\beta(1)\delta(2)$, where the greek letters denote different spins.

$^1\text{H}_2$

We go through a checklist

1. The ground state is $^1\Sigma_g^+$
 - (a) The overall wavefunction is symmetric to σ_{elec} .
 - (b) The overall wavefunction is also symmetric to i_{elec} .
2. ^1H has $I = 1/2$
 - (a) It is a fermion, the overall wavefunction must be antisymmetric wrt. nuclear exchange.
 - (b) There are 3 symmetric nuclear spin states and 1 antisymmetric nuclear spin states.

We can now generate the following table

| rotational | nuclear spin | C_2 | i_{elec} | σ_{elec} | ρ_{nuc} | overall |
|------------|--------------|-------|-------------------|------------------------|---------------------|--------------|
| even J | sym | + | + | + | +(3) | +(forbidden) |
| even J | antisym | + | + | + | -(1) | -(permitted) |
| odd J | sym | - | + | + | +(3) | -(permitted) |
| odd J | antisym | - | + | + | -(1) | +(forbidden) |

This means that rotational states with odd J values have three times the statistical weight than those with even values. This will result in the a Raman spectrum with alternately high and low peaks:

$$\begin{aligned}
 I_{\text{odd } J} &\propto 3 \times (2J + 1) \exp \left[\frac{-BJ(J+1)}{kT} \right] \\
 I_{\text{even } J} &\propto 1 \times (2J + 1) \exp \left[\frac{-BJ(J+1)}{kT} \right]
 \end{aligned}
 \tag{7.2.49}$$

where the extra factor in front is the ratio between the numbers of *ortho* (sym) and *para* (antisym) H_2 molecules. The high temperature equilibrium is intuitively 3 : 1, but we shall prove it in the next section.

 $^{14}\text{N}_2$

1. The ground state is $^1\Sigma_g^+$
 - (a) The overall wavefunction is symmetric to σ_{elec} .

(b) The overall wavefunction is also symmetric to i_{elec} .

2. ^{14}N has $I = 1$

(a) It is a boson, the overall wavefunction must be antisymmetric wrt. nuclear exchange.

(b) There are 6 symmetric nuclear spin states and 3 antisymmetric nuclear spin states.

The table is

| rotational | nuclear spin | C_2 | i_{elec} | σ_{elec} | ρ_{nuc} | overall |
|------------|--------------|-------|-------------------|------------------------|---------------------|--------------|
| even J | sym | + | + | + | +(6) | +(permitted) |
| even J | antisym | + | + | + | -(3) | -(forbidden) |
| odd J | sym | - | + | + | +(6) | -(forbidden) |
| odd J | antisym | - | + | + | -(3) | +(permitted) |

The ratio of odd and even J states are now 1:2 (ortho:para nitrogen).

$^{16}\text{O}_2$ (spin zero)

1. The ground state is $^3\Sigma_g^-$

(a) The overall wavefunction is antisymmetric to σ_{elec} .

(b) The overall wavefunction is symmetric to i_{elec} .

2. ^{16}O has $I = 0$

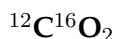
(a) It is a boson, the overall wavefunction must be antisymmetric wrt. nuclear exchange.

(b) There are 1 symmetric nuclear spin states and 0 antisymmetric nuclear spin states.

The table is thus

| rotational | nuclear spin | C_2 | i_{elec} | σ_{elec} | ρ_{nuc} | overall |
|------------|--------------|-------|-------------------|------------------------|---------------------|--------------|
| even J | sym | + | + | - | +(1) | -(forbidden) |
| odd J | sym | - | + | - | +(1) | +(permitted) |

$^{16}\text{O}_2$ can only exist in odd J states, which clearly will have an effect on the Raman spectrum ($10\tilde{B}_0, 8\tilde{B}_0, 8\tilde{B}_0, \dots$)



C lies in the centre of symmetry so is not affected by nuclear exchange, so we just need to consider the two oxygen atoms.

1. The ground state is $^1\Sigma_g^+$

(a) The overall wavefunction is symmetric to σ_{elec} .

(b) The overall wavefunction is antisymmetric to i_{elec} .

2. ^{16}O has $I = 0$ (same as above)

(a) It is a boson, the overall wavefunction must be antisymmetric wrt. nuclear exchange.

(b) There are 1 symmetric nuclear spin states and 0 antisymmetric nuclear spin states.

Therefore we have

| rotational | nuclear spin | C_2 | i_{elec} | σ_{elec} | ρ_{nuc} | overall |
|------------|--------------|-------|-------------------|------------------------|---------------------|--------------|
| even J | sym | + | + | + | +(1) | +(permitted) |
| odd J | sym | − | + | + | +(1) | −(forbidden) |

So this is clearly the opposite case as the dioxygen, in the *pure rotational Raman* spectrum, only even S lines are seen, with spacings of $6\tilde{B}_0, 8\tilde{B}_0, 8\tilde{B}_0 \dots$

The effect on the *IR spectrum* however is subtly different, as the selection rule of IR spectrum is $\Delta J = \pm 1$, we might expect the *PR* rotational fine structure to not come up at all, but this is not the case. Thinking carefully, we have been working on the assumption that all the rotational states belong to the ground state, as is the case for *pure rotational Raman*, so it has been implicit that there's a column for 'vibrational', under which it has always been symmetrical (as all ground state vibrational modes are), and that will be absorbed into the sign/parity under C_2 . So now, with the IR-active normal mode, we have alternately symmetric and antisymmetric vibrational wavefunctions, and so we can generate

| vibrational | rotational | nuclear spin | C_2 | i_{elec} | σ_{elec} | ρ_{nuc} | overall |
|-------------|------------|--------------|-------|-------------------|------------------------|---------------------|--------------|
| $v = 0$ | | | | | | | |
| sym | even J | sym | + | + | + | +(1) | +(permitted) |
| sym | odd J | sym | − | + | + | +(1) | −(forbidden) |
| $v = 1$ | | | | | | | |
| antisym | even J | sym | − | + | + | +(1) | −(forbidden) |
| antisym | odd J | sym | + | + | + | +(1) | +(permitted) |

and we can clearly see that now transitions between even and odd J are allowed, *given that* the vibrational levels also change in parity, the *rotational fine structure* will appear ‘normal’. The same argument applies for the Raman-active modes, however the rotation fine structures there are less resolved.

7.2.4.3 Consequences on rotational partition function

The symmetry factor σ_{rot} arises from the fact that *nuclear spin and rotational wavefunctions are coupled* in homonuclear diatomics (and molecules with an array of other symmetries, but it’s too complicated to go into those), so we have to evaluate the two partition functions together. For heteronuclear diatomics they are uncoupled and they can be simply computed separately:

$$q_{\text{NS}}q_{\text{rot}} = (2I_1 + 1)(2I_2 + 1) \sum_J (2J + 1) \exp\left(\frac{-BJ(J + 1)}{kT}\right) \quad (7.2.50)$$

But in a homonuclear diatomic, we have to write

$$\begin{aligned} q_{\text{NS}}q_{\text{rot}} = & g_{\text{NS, even}} \sum_{\text{even } J} (2J + 1) \exp\left(\frac{-BJ(J + 1)}{kT}\right) \\ & + g_{\text{NS, odd}} \sum_{\text{odd } J} (2J + 1) \exp\left(\frac{-BJ(J + 1)}{kT}\right) \end{aligned} \quad (7.2.51)$$

in the limit of $kT \gg B$, there are many terms contributing to the sum, and the odd and even sums are roughly equal to if all J were to be summed up, which would have given kT/B , so the odd and even sums here are both equal to $kT/2B$, and the product is

$$q_{\text{NS}}q_{\text{rot}} = (g_{\text{NS, even}} + g_{\text{NS, odd}}) \frac{kT}{2B} = g_{\text{NS}} \frac{kT}{2B} \quad (7.2.52)$$

which effectively has given us a factor of 2 in the denominator, which is the origin of the symmetry factor σ_{rot} .

7.3 CHEMICAL EQUILIBRIUM

7.3.1 Preliminaries

7.3.1.1 Chemical potential

Referring to section 6.5.5, we can write the Gibbs free energy as

$$G = \sum \mu_i N_i \quad (7.3.1)$$

This applies for all homogeneous mixtures. From the master equation we also have

$$\mu_i = \left(\frac{\partial G}{\partial n_i} \right)_{p,T,n_j} \quad (7.3.2)$$

at constant p and T we can write

$$dG = \sum_i dn_i \mu_i \quad (7.3.3)$$

Consider the following chemical reaction



we can write the change in Gibbs energy when dz amount of reaction takes place:

$$dG = +\nu_M dz \mu_M + \nu_N dz \mu_N - \nu_A dz \mu_A - \nu_B dz \mu_B \quad (7.3.5)$$

which can be rearranged to give

$$\frac{dG}{dz} = +\nu_M \mu_M + \nu_N \mu_N - \nu_A \mu_A - \nu_B \mu_B \quad (7.3.6)$$

And at equilibrium, we have $dG/dz = 0$, so

$$0 = +\nu_M \mu_M^{\text{eq}} + \nu_N \mu_N^{\text{eq}} - \nu_A \mu_A^{\text{eq}} - \nu_B \mu_B^{\text{eq}} \quad (7.3.7)$$

Switching to working with Helmholtz energy, we have

$$\mu_i = \left(\frac{\partial A}{\partial N_i} \right)_{V,T,N_j} \quad (7.3.8)$$

which can be computed by using

$$A = -NkT \ln q + kT(N \ln N - N) \quad (7.3.9)$$

so, assuming the species are non-interacting, we can write

$$\mu_i = -kT \ln \frac{q_i}{N_i} \quad (7.3.10)$$

7.3.1.2 Choice of energy zero

| Energy | Zero |
|---------------|-------------------|
| Translational | Already zero |
| Rotational | Already zero |
| Vibrational | Bottom of well |
| Electronic | Dissociated atoms |

So we can choose to write the Helmholtz energy as

$$A = -NkT \ln q + kT(N \ln N - N) + N\epsilon^0 \quad (7.3.11)$$

where

$$q = q_{\text{trans}} q_{\text{rot}} q_{\text{vib}} q_{\text{el}} \quad (7.3.12)$$

and q_{vib} refers to bottom of well and q_{el} takes electronic ground state as energy, zero, and the referral to dissociated atoms is done in the *offset term* $N\epsilon^0$.³ For species i in an *ideal* mixture this is then

$$\mu_i = -kT \ln \frac{q_i}{N_i} + \epsilon^{0,i} \quad (7.3.13)$$

7.3.2 Equilibrium constants

7.3.2.1 Calculation of K_c

Considering the same chemical reaction



³This is justified as $N \ln \sum_i g_i \exp(-\epsilon^i/kT) = N\epsilon^0 + N \ln \sum_i g_i \exp(-(\epsilon^i - \epsilon^0)/kT)$.

with

$$0 = +\nu_M \mu_M^{\text{eq}} + \nu_N \mu_N^{\text{eq}} - \nu_A \mu_A^{\text{eq}} - \nu_B \mu_B^{\text{eq}} \quad (7.3.15)$$

where we will omit the 'eq' subscript and take it as implicit from now.

Only q_{trans} contains a volume term, so we can write

$$q = fV \quad (7.3.16)$$

where

$$f = \frac{q_{\text{trans}}}{V} q_{\text{rot}} q_{\text{vib}} q_{\text{el}} = \left(\frac{2\pi m k T}{h^2} \right)^{3/2} q_{\text{rot}} q_{\text{vib}} q_{\text{el}} \quad (7.3.17)$$

so

$$\mu_i = -kT \ln \frac{f_i V}{N_i} + \epsilon^{0,i} = -kT \ln \frac{f_i}{c_i} + \epsilon^{0,i} \quad (7.3.18)$$

Clearly, sticking to the SI units, the units of c_i is molecules / m³. Substituting this into eq. 7.3.15, we get

$$0 = +\nu_M \left(-kT \ln \frac{f_M}{c_M} + \epsilon^{0,M} \right) + \nu_N \left(-kT \ln \frac{f_N}{c_N} + \epsilon^{0,N} \right) - \nu_A \left(-kT \ln \frac{f_A}{c_A} + \epsilon^{0,A} \right) - \nu_B \left(-kT \ln \frac{f_B}{c_B} + \epsilon^{0,B} \right) \quad (7.3.19)$$

Rearrangement gives

$$kT \ln \frac{c_M^{\nu_M} c_N^{\nu_N}}{c_A^{\nu_A} c_B^{\nu_B}} = kT \ln \frac{f_M^{\nu_M} f_N^{\nu_N}}{f_A^{\nu_A} f_B^{\nu_B}} - \Delta \epsilon_0 \quad (7.3.20)$$

where $\Delta \epsilon_0$ is the energy of products minus reactants. Cleaning up further we have

$$K'_c = \frac{c_M^{\nu_M} c_N^{\nu_N}}{c_A^{\nu_A} c_B^{\nu_B}} = \frac{f_M^{\nu_M} f_N^{\nu_N}}{f_A^{\nu_A} f_B^{\nu_B}} \exp(-\Delta \epsilon_0 / kT) \quad (7.3.21)$$

Recalling the proper thermodynamic K_c is

$$K_c = \frac{\alpha_M^{\nu_M} \alpha_N^{\nu_N}}{\alpha_A^{\nu_A} \alpha_B^{\nu_B}} \quad (7.3.22)$$

α_i in an *ideal* aqueous solution is

$$\alpha_i = \frac{c_i}{c^\circ} \quad (7.3.23)$$

So we can obtain the 'actual' K_c easily

$$K_c = K'_c \left(\frac{1}{c^\circ} \right)^{\Delta \nu} \quad (7.3.24)$$

The reference state c° is *always* given in units of molecules / m³, as this expression arise from statistical mechanics. So for a final K_c referred to the usual standard state of 1 mol / dm³, c° is chosen to be $1000N_A$.

So we arrive at

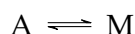
Theorem 7.3.2.1 (Expression of K_c). The full expression of K_c is

$$K_c = \frac{f_M^{\nu_M} f_N^{\nu_N}}{f_A^{\nu_A} f_B^{\nu_B}} \left(\frac{1}{c^\circ} \right)^{\Delta\nu} \exp(-\Delta\epsilon_0/kT) \quad (7.3.25)$$

where $\Delta\epsilon_0$ is the energy change, at absolute zero, per molecule, and K_c is dimensionless.

Interpretation

Considering a simple reaction



We can write the K_c as

$$K_c = \frac{f_M(T)}{f_A(T)} \exp(-\Delta\epsilon_0/kT) \quad (7.3.26)$$

where the temperature dependence is made explicit, at low temperatures, only ground states are accessible and the exponential term is large, so it predominates, the equilibrium constant is less than one. At high temperatures, the ratio f_M/f_A depends on the density of states of the two species - if the M states are denser than the A states, the ratio predominates the exponential, and the reaction will be entropically driven to the right.

7.3.2.2 Calculation of K_p

If we start with eq. 6.5.44, writing

$$\Delta_r G^\circ = \nu_M \mu_M^\circ + \nu_N \mu_N^\circ - \nu_A \mu_A^\circ - \nu_B \mu_B^\circ \quad (7.3.27)$$

where

$$\mu_i^\circ = N_A \left(-kT \ln \frac{q_i^\circ}{N_A} + \epsilon^{0,i} \right) = -RT \ln \frac{q_i^\circ}{N_A} + \epsilon_m^{0,i} \quad (7.3.28)$$

in which q° is the molar partition function in the standard state, which only concerns the volume, which will be the standard molar volume V_m° ⁴, so we can write

$$\begin{aligned} q^\circ &= q_{\text{trans}} q_{\text{rot}} q_{\text{vib}} q_{\text{el}} \\ &= \left(\frac{2\pi m k T}{h^2} \right)^{3/2} V_m^\circ q_{\text{rot}} q_{\text{vib}} q_{\text{el}} \\ &= \left(\frac{2\pi m k T}{h^2} \right)^{3/2} \left(\frac{RT}{p^\circ} \right) q_{\text{rot}} q_{\text{vib}} q_{\text{el}} \end{aligned} \quad (7.3.29)$$

Substituting eq. 7.3.28 into eq. 7.3.27 we can have the following result

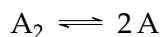
Theorem 7.3.2.2 (Expression of K_p). The equilibrium constant is given by

$$K_p = \frac{(q_M^\circ)^{\nu_M} (q_N^\circ)^{\nu_N}}{(q_A^\circ)^{\nu_A} (q_B^\circ)^{\nu_B}} (N_A)^{-\Delta\nu} \exp\left(-\frac{\Delta\epsilon_0}{kT}\right) \quad (7.3.30)$$

7.3.3 Examples

7.3.3.1 Dissociation of a diatomic

We want to calculate the K_c of the following reaction



We note that $\Delta\nu = 1$, $\Delta\epsilon_0 = D_e$, so

$$K_c = \frac{f_A^2}{f_{A_2}} \frac{1}{c^\circ} \exp(-D_e/kT) \quad (7.3.31)$$

The key thing to remember is just that σ for A_2 is 2, and also that the electronic ground state degeneracies might differ.

Different isotopes can also give different K_c , now consider the two reactions



At first sight we don't expect the rate constants to be different, but they the first is two times bigger than the second, even after taking the difference in masses into account. The

⁴Remember that temperature is not part of the standard state, so temperature is still an independent variable.

reason will be clear if we write out the expressions for the equilibrium constants:

$$\begin{aligned} K_c &= \frac{f_{\text{HD}}}{f_{\text{H}}f_{\text{D}}} \left(\frac{1}{c^\circ} \right)^{-1} \exp(-\Delta\epsilon_0/kT) \\ K_c &= \frac{f_{\text{H}_2}}{f_{\text{H}}^2} \left(\frac{1}{c^\circ} \right)^{-1} \exp(-\Delta\epsilon_0/kT) \end{aligned} \quad (7.3.33)$$

Just focussing on the electronic and rotational parts of the wavefunction, as the vibrational and translational (difference in masses already accounted for) are the same, we can write

$$K_c \propto \frac{g_{0,\text{AB}}}{g_{0,\text{A}}g_{0,\text{B}}} \frac{1}{\sigma} \quad (7.3.34)$$

For the first reaction, the first bit will return $\frac{1}{4}$, and this can be interpreted as the fact that only when spins of the two atoms are antiparallel (singlet) can the two atoms form a bond, which means only one in four spin configurations can lead to reaction, the other three being the triplet state.

For the second reaction, the first bit is also $\frac{1}{4}$, but the second is $\frac{1}{2}$, which can be interpreted as that half of all rotational-nuclear states are disallowed, as the two are coupled in a homonuclear diatomic, see section 7.2.4.2.

7.3.3.2 Ionisation of atoms

The equation is



We again note that $\Delta\nu = 1$, $\Delta\epsilon_0 = E_1$, the ionisation energy (per molecule of course, if per mole divide by RT instead), so

$$K_c = \frac{f_{\text{e}}f_{\text{M}^+}}{f_{\text{M}}} \frac{1}{c^\circ} \exp(-E_1/kT) \quad (7.3.35)$$

The masses of M and M^+ cancel out, electron has $g_0 = 2$ as it has spin one half, so we can write

$$K_c = \left(\frac{2\pi m_{\text{e}}kT}{h^2} \right)^{3/2} \frac{2g_{0,\text{M}^+}}{g_{0,\text{M}}} \frac{1}{c^\circ} \exp -E_1/kT \quad (7.3.36)$$

cleaning up we have the

Theorem 7.3.3.1 (The Saha equation). The Saha equation is used in the spectral classification of stars, which can be treated as gases at high enough temperatures, and it's given by

$$K_c = \frac{2}{\Lambda^3} \frac{g_+}{g_0} \exp(-E_1/kT) \quad (7.3.37)$$

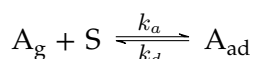
where

$$\Lambda = \sqrt{\frac{h^2}{2\pi m_e kT}} \quad (7.3.38)$$

is the thermal wavelength of an electron, and g_+ is the ground state degeneracy of the ion and g_0 of the atom.

7.3.3.3 Adsorption of a gas

The general equation for adsorption is



where A_g is the adsorbate molecule, S are sites on a surface, and A_{ad} is the adsorbed molecule

Theorem 7.3.3.2 (Langmuir isotherm). The Langmuir isotherm governs the proportion of gas that's adsorbed at equilibrium and it's given by

$$\theta = \frac{K_{ads}p}{1 + K_{ads}p} \quad (7.3.39)$$

Proof. The rate of adsorption is proportional to the amount of available sites and the gas pressure:

$$R_{adsorption} = k'_a[S][A_g] = k_a \times (M - N) \times p \quad (7.3.40)$$

where k_a absorbs an arbitrary conversion factor from $[A_g]$ to p .

The rate of desorption is only proportional to the amount of occupied sites:

$$R_{desorption} = k'_d[A_{ad}] = k_d \times N \quad (7.3.41)$$

At equilibrium the two must be equal, so we can write

$$K_{ads} = \frac{k_a}{k_d} = \frac{N}{M - N} \frac{1}{p} = \frac{\theta}{1 - \theta} \frac{1}{p} \quad (7.3.42)$$

Which rearranges easily to give the Langmuir isotherm. □

Monatomic gas

Now we wish to find an expression of K_{ads} for the simplest case of a monatomic gas. We note that at equilibrium

$$\mu(\text{ads}) = \mu(\text{g}) \quad (7.3.43)$$

for the free adsorbate gas, we further assume that it's a monatomic gas, so

$$\mu(\text{ads}) = kT \ln \left(\frac{q_{\text{trans}}}{N_{\text{gas}}} \right) + \epsilon_{\text{free}}^0 = -kT \ln \left(\frac{kT}{p\Lambda^3} \right) + \epsilon_{\text{free}}^0 \quad (7.3.44)$$

⁵ For the adsorbed atoms it's less straightforward, we first write the Helmholtz energy for the system as

$$A = -kT \ln Q_N \quad (7.3.45)$$

We then note that

$$Q_N = \sum_m \exp(-E_m/kT) \quad (7.3.46)$$

and we assume that all arrangements of the atoms have the same energy *i.e.*, the sites are equivalent and non-interacting, so Q_N is simply the number of ways to arrange N atoms in M sites where $M \gg N$:

$$Q_N = \frac{M!}{N!(M-N)!} \quad (7.3.47)$$

so

$$A = kT[-M \ln M + N \ln N + (M-N) \ln(M-N)] \quad (7.3.48)$$

and by definition

$$\mu = \left(\frac{\partial A}{\partial N} \right)_{T,V} \quad (7.3.49)$$

with the offset correction this computes to

$$\mu(\text{ads}) = kT \ln \left(\frac{N}{M-N} \right) + \epsilon_{\text{ads}}^0 = kT \ln \left(\frac{\theta}{1-\theta} \right) + \epsilon_{\text{ads}}^0 \quad (7.3.50)$$

equating the two chemical potentials we have

$$-kT \ln \left(\frac{kT}{p\Lambda^3} \right) + \epsilon_{\text{free}}^0 = kT \ln \left(\frac{\theta}{1-\theta} \right) + \epsilon_{\text{ads}}^0 \quad (7.3.51)$$

which rearranges to

$$\left(\frac{\theta}{1-\theta} \right) \frac{1}{p} = \frac{\Lambda^3}{kT} \exp \left(\frac{-\Delta\epsilon^0}{kT} \right) \quad (7.3.52)$$

⁵We have ignored the electronic partition functions here as they cancel out. They should generally be included.

the LHS of the equation can be identified as K_{ads} , the expression for θ is then

$$\theta = \frac{bp}{1 + bp} \quad (7.3.53)$$

where

$$b = \frac{\Lambda^3}{kT} \exp\left(\frac{-\Delta\epsilon^0}{kT}\right) \quad (7.3.54)$$

and

$$\Lambda = \sqrt{\frac{h^2}{2\pi mkT}} \quad (7.3.55)$$

Theorem 7.3.3.3 (Expression of K_{ads} for monatomic gas). The adsorption equilibrium constant of a monatomic gas is given by

$$K_{\text{ads}} = \frac{\Lambda^3}{kT} \exp(-\Delta\epsilon^0/kT) \quad (7.3.56)$$

with the fraction of adsorbed atoms given by

$$\theta = \frac{K_{\text{ads}}p}{1 + K_{\text{ads}}p} \quad (7.3.57)$$

Non-dissociative molecular adsorption

This is almost exactly the same as the monatomic gas except now there are internal contributors of partition function:

$$Q_{N,\text{ads}} = q_{\text{ads}}^N \frac{M!}{N!(M-N)!} \quad (7.3.58)$$

where $q_{\text{ads}} = q_{\text{vib}}q_{\text{rot}}$ is the internal degrees of freedom. We've omitted the electronic partition functions again as they'll cancel out. The chemical potential is thus

$$\mu(\text{ads}) = -kT \ln q_{\text{ads}} + kT \ln \left(\frac{\theta}{1 - \theta} \right) + \epsilon_{\text{ads}}^0 \quad (7.3.59)$$

where $\theta = N/M$, the *coverage factor*.

For the free gas,

$$\mu(\text{g}) = -kT \ln \left(\frac{q}{N} \right) + \epsilon_{\text{free}}^0 = -kT \ln \left(\frac{kT}{p\Lambda^3} q_{\text{free}} \right) + \epsilon_{\text{free}}^0 \quad (7.3.60)$$

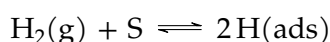
where again q_{free} is the internal degrees of freedom, omitting electronic. We go through the same steps to obtain

Theorem 7.3.3.4 (Non-dissociative adsorption). The equilibrium constant is given as

$$K_{\text{ads}} = \frac{q_{\text{ads}}}{q_{\text{free}}} \frac{\Lambda^3}{kT} \exp(-\Delta\epsilon^0/kT) \quad (7.3.61)$$

Dissociative molecular adsorption

Now we consider the reaction



For N H_2 gas molecules, there are M sites for H *atoms*, the number of ways to fit the molecules onto the sites is, in the limit that $M \gg N$,

$${}^{M/2}C_N = \frac{(M/2)!}{N!(M/2 - N)!} \quad (7.3.62)$$

where the the hydrogen molecules have effective access to $M/2$ sites. The partition function of the system is then

$$Q_N = \frac{(M/2)!}{N!(M/2 - N)!} q_{\text{H}}^{2N} \quad (7.3.63)$$

where all the hydrogen atoms are distinguishable, and the partition function just includes the electronic partition function. So the chemical potential is

$$\mu(\text{ads}) = -2kT \ln q_{\text{H}} + kT \ln \left(\frac{\theta}{1 - \theta} \right) + \epsilon_{\text{ads}}^0 \quad (7.3.64)$$

For the free gas it is the same as before,

$$\mu(\text{g}) = -kT \ln \left(\frac{kT}{p\Lambda^3} q_{\text{H}_2} \right) + \epsilon_{\text{free}}^0 \quad (7.3.65)$$

At equilibrium,

$$\mu(\text{g}) = 2\mu(\text{ads}) \quad (7.3.66)$$

and so we can write

$$\begin{aligned} -kT \ln \left(\frac{kT}{p\Lambda^3} q_{\text{H}_2} \right) + \epsilon_{\text{free}}^0 &= 2 \left[-2kT \ln q_{\text{H}} + kT \ln \left(\frac{\theta}{1 - \theta} \right) + \epsilon_{\text{ads}}^0 \right] \\ -\ln \left(\frac{kT}{p\Lambda^3} q_{\text{H}_2} \right) &= -2 \ln q_{\text{H}}^2 + \ln \left(\frac{\theta}{1 - \theta} \right)^2 + \Delta\epsilon^0/kT \end{aligned} \quad (7.3.67)$$

on hold, the derivation is not satisfactory: do we assume adjacent sites? kinetic argument assuming adjacency obtains the square roots but statmech arguments can only get square

roots if we use M choose $2N$, which does not assume adjacency.

7.4 RATE CONSTANTS

7.4.1 Transition state theory

7.4.1.1 The theory

The theory is essentially the assumption of the following kinetic scheme



The equilibrium constant for formation of transition state, K^* , is given by

$$K^* = \frac{\alpha_{TS}}{\alpha_A \alpha_{BC}} = \frac{[TS]}{[A][BC]} c^{\circ'} \quad (7.4.1)$$

where $c^{\circ'}$ is the reference concentration, the units of which does not necessarily need to be in molecules / m³, because this does not arise from statistical mechanics. However this does arise from classical thermodynamics so $c^{\circ'}$ must be 1 mol dm⁻³, expressed in the units used to measure the concentrations.

The rate at which the transition state breaks down is assumed to be simply first-order:

$$r = k_{1st}[TS] \quad (7.4.2)$$

where from one equation before we can substitute in the expression for [TS]:

$$r = k_{1st} K^* [A][BC] \frac{1}{c^{\circ'}} \quad (7.4.3)$$

The overall rate of reaction is the rate of production of the products, which is the rate at which transition state gives products. As overall the reaction is bimolecular, it should be

$$r = k_{2nd}[A][BC] \quad (7.4.4)$$

equating the two expressions we have

$$k_{2nd} = \frac{1}{c^{\circ'}} k_{1st} K^* \quad (7.4.5)$$

7.4.1.2 Calculation of rate constant

To calculate k_{2nd} , we just need to know k_{1st} and K^* . For K^* , we can invoke theorem 7.3.2.1 and write

$$K^* = \frac{f_{TS}}{f_A f_{BC}} \left(\frac{1}{c^\circ} \right)^{-1} \exp(-\Delta\epsilon_0^\ddagger/kT) \quad (7.4.6)$$

where c° is technically arbitrary, expressed in the units of molecule / m³, however, as we're using it in a thermodynamic expression, we should set it to $1000N_A$ so that we always refer to the thermodynamic standard state.

In dealing with f_{TS} we need to realise one of the vibrational modes, usually the asymmetric stretch, will lead to the reaction, as the TS is at energy maxima, that one vibrational mode will have no restoring force. We write the q_{vib} due to that mode q^\ddagger :

$$K^* = \frac{q^\ddagger f'_{TS}}{f_A f_{BC}} \left(\frac{1}{c^\circ} \right)^{-1} \exp(-\Delta\epsilon_0^\ddagger/kT) \quad (7.4.7)$$

q^\ddagger can be computed as

$$q^\ddagger = \frac{\exp(-h\nu^\ddagger/2kT)}{1 - \exp(-h\nu^\ddagger/kT)} \quad (7.4.8)$$

The frequency associated with this vibrational mode is likely to be very low as it is positively related to the spring constant (restoring force), in this case it is zero. Therefore under the limit that $h\nu^\ddagger/kT \ll 1$, we can approximate the above expression to

$$q^\ddagger \approx \frac{kT}{h\nu^\ddagger} \quad (7.4.9)$$

So in full,

$$K^* = \frac{kT}{h\nu^\ddagger} \frac{f'_{TS}}{f_A f_{BC}} \left(\frac{1}{c^\circ} \right)^{-1} \exp(-\Delta\epsilon_0^\ddagger/kT) \quad (7.4.10)$$

Meanwhile, k_{1st} is just the rate at which the transition state breaks down, which can be identified with ν^\ddagger , so we can write k_{2nd} as

$$\begin{aligned} k_{2nd} &= \frac{1}{c^{\circ'}} k_{1st} K^* \\ &= \frac{1}{c^{\circ'}} \nu^\ddagger \frac{kT}{h\nu^\ddagger} \frac{f'_{TS}}{f_A f_{BC}} \left(\frac{1}{c^\circ} \right)^{-1} \exp(-\Delta\epsilon_0^\ddagger/kT) \end{aligned} \quad (7.4.11)$$

now before proceeding to simplify this further we must clarify what the reference concentrations are here:

- $c^{\circ'}$ is the thermodynamic reference concentration, expressed in whatever units that are used to measure the concentration in, referred to the thermodynamic standard

state of 1 mol dm^{-3} . Here we assume the concentration is measured in units of mol dm^{-3} , so simply $c^\circ = 1$

- c° is the statmech reference concentration, expressed in units of $1 \text{ molecule m}^{-3}$. For the expression to be coherent we must set $c^\circ = 1000N_A$

Therefore we have

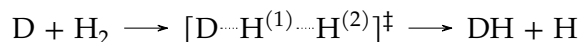
Theorem 7.4.1.1 (Rate constant). $k_{2\text{nd}}$ in units of $\text{M}^{-1} \text{ s}^{-1}$ is given as

$$k_{2\text{nd}} = \frac{kTc^\circ}{h} \frac{f'_{\text{TS}}}{f_A f_{\text{BC}}} \exp(-\Delta\epsilon_0^\ddagger/kT) \quad (7.4.12)$$

7.4.2 Applications of transition state theory

7.4.2.1 Rate constant calculation

For the very few, very simplest reactions for which we have the complete PES, we can calculate the rate constant for them. Now for the reaction



we have the data:

- Reactants

- $r_{\text{H}_2} = 0.741 \text{ \AA}$
- $\tilde{\nu}_{\text{H}_2} = 4401 \text{ cm}^{-1}$

- Transition state

- $r_{\text{D}-\text{H}^{(1)}} = 0.929 \text{ \AA}$
- $r_{\text{H}^{(1)}-\text{H}^{(2)}} = 0.929 \text{ \AA}$
- $\tilde{\nu}_{\text{sym}} = 1708 \text{ cm}^{-1}$
- $\tilde{\nu}_{\text{bend}} = 861 \text{ cm}^{-1}$, doubly degenerate
- $\Delta\epsilon_0^\ddagger = 40.3 \text{ kJ mol}^{-1}$

We wish to find the rate constant $k_{2\text{nd}}$ at 1000 K. Now breaking our calculation into parts, we shall compute f'_{TS} first:

$$f'_{\text{TS}} = f_{\text{trans}} \frac{q_{\text{vib}}}{q^\ddagger} q_{\text{rot}} q_{\text{el}} \quad (7.4.13)$$

One by one, f_{trans} is simply

$$f_{\text{trans}} = \left(\frac{2\pi m_{\text{H}_2\text{D}} kT}{h^2} \right)^{3/2} = 4.754 \times 10^{31} \quad (7.4.14)$$

the vibrational partition function is

$$\frac{q_{\text{vib}}}{q^\ddagger} = q_{\text{sym}} q_{\text{bend}}^2 = 0.1838 \quad (7.4.15)$$

note we were not given q^\ddagger , which would have been the antisymmetric stretch. The next is q_{rot} :

$$q_{\text{rot}} = \frac{T}{\sigma \theta_{\text{rot}}} = \frac{2IkT}{\sigma \hbar^2} = \frac{2\mu r^2 kT}{\sigma \hbar^2} \quad (7.4.16)$$

where the reduced mass for a linear triatomic $m_1 - m_2 - m_3$ is given by

$$\mu = m_1 r_1^2 + m_3 r_2^2 - \frac{(m_1 r_1 - m_3 r_2)^2}{m_1 + m_2 + m_3} \quad (7.4.17)$$

so at 1000 K $q_{\text{rot}} = 97.85$.

Now, for q_{el} , we note that D has the electronic ground state of $^2S_{1/2}$ and H_2 has $^1\Sigma^+$, so the system has overall one unpaired electron, so the spin multiplicity of 2 needs to be preserved throughout the reaction, so $q_{\text{el}}^\ddagger = 2$.

All in, $f'_{\text{TS}} = 1.708 \times 10^{33}$.

The reactant partition functions are

$$\begin{aligned} f_{\text{D}} &= 3.362 \times 10^{31} \\ f_{\text{H}_2} &= 4.018 \times 10^{30} \end{aligned} \quad (7.4.18)$$

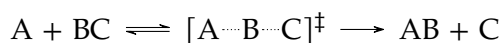
remembering that q_{el} for D is 2. Putting this into theorem 7.4.1.1, we have

$$k_{2\text{nd}} = 1.245 \times 10^9 \text{ M}^{-1} \text{ s}^{-1} \quad (7.4.19)$$

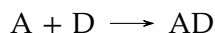
7.4.2.2 Steric factor

The theory is rarely used to directly compute the rate constants as detailed PES is required, and it is often used to derive qualitative results like this.

We would like to find out the steric factor for the reaction



where the transition state is assumed to be linear. The steric factor is defined as the ratio of rates between a sterically restricted reaction and a sterically unrestricted reaction, such as below



As the different contributors to the overall partition function different in many orders of magnitude, we can write, say

$$f_{BC} = [\text{trans}]^3 [\text{rot}]^2 [\text{vib}] \quad (7.4.20)$$

where each degree of freedom contribute to one power, so trans is always cubed, rot is to the power of 2 if it's a linear molecule and 3 if it's non-linear, and vib is $3N - 5$ for linear and $3N - 6$ for non-linear, but remember to remove 1 for f'_{TS} for the mode that leads to reaction. So for A,

$$f_A = [\text{trans}]^3 \quad (7.4.21)$$

and for the transition state,

$$f'_{TS} = [\text{trans}]^3 [\text{rot}]^2 [\text{vib}]^3 \quad (7.4.22)$$

For the structureless reaction, we can write

$$f_D = [\text{trans}]^3 \quad (7.4.23)$$

and

$$f'_{TS} = [\text{trans}]^3 [\text{rot}]^2 \quad (7.4.24)$$

With the activation energy being assumed equal, the ratio is

$$p = \frac{k_{2nd}(A + BC)}{k_{2nd}(A + D)} = \frac{[\text{vib}]^2}{[\text{rot}]^2} \quad (7.4.25)$$

Now to give these numerical values, we take the typical vibrational frequency at 800 cm^{-1} , which gives $[\text{vib}] = 0.15$ per vibrational freedom, and for the typical linear diatomic rotational frequency at 0.5 cm^{-1} , we have $[\text{rot}]^2 = 414$, so per rotational freedom is about 20 *i.e.*, $[\text{rot}] = 20$.⁶ So in this case $p = 5 \times 10^{-5}$, which is very small.

If the transition state is bent, we need to write the transition state partition function as

$$f'_{TS} = [\text{trans}]^3 [\text{rot}]^3 [\text{vib}]^2 \quad (7.4.26)$$

⁶In computing the energy levels we have assumed that the diatomic are particles on a sphere, which assumes two degrees of rotational freedom already.

so we end up with

$$p = \frac{[\text{vib}]}{[\text{rot}]} \approx 7 \times 10^{-3} \quad (7.4.27)$$

which is larger because the a range of angles are now possible.

7.4.2.3 Kinetic isotope effects

The electronic structure *i.e.*, the Morse potential of the molecule cannot be affected by substitution by an isotope as the nuclear charges remain the same; and also the moments of inertia, hence the rotational wavefunctions, must remain relatively unaffected for large molecules. However the vibrational levels are affected as those are more localised (normal modes are practically localised in big molecules). As $\nu \propto \frac{1}{\mu}$, the ratio of the frequencies are

$$\frac{\nu_A}{\nu_B} = \sqrt{\frac{\mu_B}{\mu_A}} \quad (7.4.28)$$

Suppose in a reaction we have a large molecule (think proteins) where a bond to a hydrogen atom is broken, if we substitute that hydrogen with deuterium, the only thing different will be the vibrational partition function, where

$$\begin{aligned} q_{\text{vib,H}} &= \frac{\exp(-h\nu_{\text{H}}/2kT)}{1 - \exp(-h\nu_{\text{H}}/2kT)} \\ q_{\text{vib,D}} &= \frac{\exp(-h\nu_{\text{D}}/2kT)}{1 - \exp(-h\nu_{\text{D}}/2kT)} \end{aligned} \quad (7.4.29)$$

in the limit of $h\nu \gg kT$, we can write

$$\begin{aligned} q_{\text{vib,H}} &\approx \exp(-h\nu_{\text{H}}/2kT) \\ q_{\text{vib,D}} &\approx \exp(-h\nu_{\text{D}}/2kT) \end{aligned} \quad (7.4.30)$$

By making these approximations we are assuming that only the ground state vibrational wavefunctions contribute, which can be seen by the fact that the zero point energy term shows up in the partition function. We should further note that $\Delta\epsilon_0^\ddagger$ is not affected by the isotope substitution, so the ratio of the rate constants can simplify to

$$\frac{k_{2\text{nd}}(\text{H})}{k_{2\text{nd}}(\text{D})} = \exp\left(\frac{-(\text{ZPE}_{\text{D}} - \text{ZPE}_{\text{H}})}{kT}\right) \quad (7.4.31)$$

The rate constants differ by a factor which depends on the difference of the vibrational zero point energies of the two species. Explained in terms of the Arrhenius equation, we can see that

$$\begin{aligned} k_{2\text{nd}}(\text{H}) &= A \exp(-E_{a,\text{H}}/RT) \\ k_{2\text{nd}}(\text{D}) &= A \exp(-E_{a,\text{D}}/RT) \end{aligned} \quad (7.4.32)$$

the difference in activation energy is just the difference in the vibrational ground states.

7.4.3 Thermodynamic formulation of transition state theory

The thermodynamic formulation can be used to calculate the parameters of the transition state theory such as activation energy from existing thermodynamic data, which provides an explanatory rather than predictive model.

7.4.3.1 Transition Gibbs energy

We recall eq. 7.4.10, which we now make the definition

$$K^* = \underbrace{\frac{kT}{h\nu^\ddagger}}_{q^\ddagger} \underbrace{\frac{f'_{\text{TS}}}{f_{\text{A}}f_{\text{BC}}} \left(\frac{1}{c^\circ}\right)^{-1} \exp(-\Delta\epsilon_0^\ddagger/kT)}_{K^\ddagger} \quad (7.4.33)$$

and we also recall theorem 7.4.1.1 which says

$$k_{2\text{nd}} = \frac{1}{c^\circ} k_{1\text{st}} K^* \quad (7.4.34)$$

where $k_{1\text{st}} = \nu^\ddagger$, putting the two together we have

$$k_{2\text{nd}} = \frac{1}{c^\circ} \frac{kT}{h} K^\ddagger \quad (7.4.35)$$

We can write

$$\Delta_r G^{\circ,\ddagger} = -RT \ln K^\ddagger \quad (7.4.36)$$

here K^\ddagger and not K^* is used because $\Delta_r G^{\circ,\ddagger}$ is the change in Gibbs energy, under standard conditions, when the transition state is formed from the reactants - if we include q^\ddagger here we include also the part of the transition state that gives reaction, which is not technically part of the equilibrium and should not be involved in the ΔG . **I'm not really sure about this** Now this is derived from classical thermodynamics which means K^\ddagger must be referred to 1 mol dm^{-3} , so this means c° in eq. 7.4.33 must be set to $1000N_A$, and to refer the standard thermodynamic state we also need to set c° in eq. 7.4.35 to 1, so we have our expression for the reformulated expression for the rate constant in solution, in units of $\text{M}^{-1} \text{ s}^{-1}$:

$$k_{2\text{nd}} = \frac{kT}{h} \exp\left(\frac{-\Delta_r G^{\circ,\ddagger}}{RT}\right) \quad (7.4.37)$$

where $\Delta_r G^{\circ,\ddagger}$ is derived from K^\ddagger referred to 1 mol dm^{-3} , as discussed above.

The gas-phase expression is more straightforward as we there's no separate statmech and thermodynamic definitions for the standard state, which is just $p^\circ = 1 \times 10^5 \text{ Pa}$. We did not derive the transition state theory expression for gas-phase reactions but the result is, in units of $\text{Pa}^{-1} \text{ s}^{-1}$:

$$k_{2\text{nd}} = \left(\frac{1}{p^\circ} \right) \frac{kT}{h} \exp \left(\frac{-\Delta_r G^{\circ,\ddagger}}{RT} \right) \quad (7.4.38)$$

These two equations are collectively known as the Eyring equation:

Theorem 7.4.3.1 (Eyring equation). The *Eyring equation* relates the rate constant of a chemical reaction with temperature and Gibbs free energy change of activation:

$$\begin{aligned} \text{Solution-phase: } k_{2\text{nd}} &= \frac{kT}{h} \exp \left(\frac{-\Delta_r G^{\circ,\ddagger}}{RT} \right) \\ \text{Gas-phase: } k_{2\text{nd}} &= \left(\frac{1}{p^\circ} \right) \frac{kT}{h} \exp \left(\frac{-\Delta_r G^{\circ,\ddagger}}{RT} \right) \end{aligned} \quad (7.4.39)$$

where $p^\circ = 1 \times 10^5 \text{ Pa}$.

We remember that the standard Gibbs energy change can be written as

$$\Delta_r G^{\circ,\ddagger} = \Delta_r H^{\circ,\ddagger} - T \Delta_r S^{\circ,\ddagger} \quad (7.4.40)$$

so

Theorem 7.4.3.2 ($k_{2\text{nd}}$ in terms of thermodynamic parameters). For a reaction in solution, the rate constant can be expressed, in units of $\text{M}^{-1} \text{ s}^{-1}$,

$$k_{2\text{nd}} = \frac{kT}{h} \exp \left(\frac{\Delta_r S^{\circ,\ddagger}}{R} \right) \exp \left(\frac{-\Delta_r H^{\circ,\ddagger}}{RT} \right) \quad (7.4.41)$$

7.4.3.2 Relationship to the Arrhenius equation

The Arrhenius equation

$$k_{2\text{nd}} = A \exp \left(\frac{-E_a}{RT} \right) \quad (7.4.42)$$

looks very much like eq. 7.4.41. But we need to be careful about the definition of activation energy from the Arrhenius equation:

$$\begin{aligned}\left(\frac{\partial \ln k_{2\text{nd}}}{\partial(1/T)}\right)_V &= -\frac{E_a}{R} \\ \left(\frac{\partial \ln k_{2\text{nd}}}{\partial T}\right)_V &= \frac{E_a}{RT^2}\end{aligned}\quad (7.4.43)$$

or basically,

$$E_a = RT^2 \left(\frac{\partial \ln k_{2\text{nd}}}{\partial T}\right)_V \quad (7.4.44)$$

and so from the expression of $k_{2\text{nd}}$ in eq. 7.4.35 we have

$$\left(\frac{\partial \ln k_{2\text{nd}}}{\partial T}\right)_V = \frac{1}{T} + \left(\frac{\partial \ln K^\ddagger}{\partial T}\right)_V \quad (7.4.45)$$

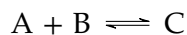
So for a reaction in solution phase,

$$E_a = \Delta_r H^{\circ,\ddagger} + RT \quad (7.4.46)$$

In gas phase, we need to use the van't Hoff isochore with K_p :

$$\frac{\partial \ln K_p}{\partial T} = \frac{\Delta_r H^\circ}{RT^2} \quad (7.4.47)$$

The K_p for the example reaction



is

$$K_p = \frac{\alpha_C}{\alpha_A \alpha_B} = \frac{p^\circ p_C}{p_A p_B} \quad (7.4.48)$$

where p° refers to 1 bar measured in Pascals. The conversion between partial pressures and concentrations is easy:

$$p_i = c_i kT \quad (7.4.49)$$

So we can express K_p in terms of K_c , for this reaction ((p°/c°) will be in different powers for different reactions.)

$$K_p = \left(\frac{c_C}{c_A c_B}\right) \frac{p^\circ}{kT} = K_c \frac{p^\circ}{c^\circ kT} \quad (7.4.50)$$

where, explicitly $p^\circ = 1 \times 10^5 \text{ Pa}$ and $c^\circ = 1000 N_A \text{ molecules m}^{-3}$, and the K_p this yields is correctly referred to 1 bar.

Substituting this expression of K_p^\ddagger into the van't Hoff isochore we have

$$\begin{aligned}\frac{\partial}{\partial T} \left(\ln K_c - \ln T + \ln \frac{p^\circ}{kc^\circ} \right) &= \frac{\Delta_r H^\circ}{RT^2} \\ \frac{\partial \ln K_c}{\partial T} &= \frac{\Delta_r H^\circ + RT}{RT^2}\end{aligned}\quad (7.4.51)$$

And we know from eq. 7.4.44 that

$$E_a = RT^2 \left(\frac{\partial \ln k_{2\text{nd}}}{\partial T} \right)_V \quad (7.4.52)$$

so we can use the result from eq. 7.4.45⁷ to write

$$\begin{aligned}E_a &= RT^2 \left[\frac{1}{T} + \left(\frac{\partial \ln K^\ddagger}{\partial T} \right)_V \right] \\ &= RT^2 \left[\frac{1}{T} + \frac{\Delta_r H^{\circ,\ddagger} + RT}{RT^2} \right] \\ &= \Delta_r H^{\circ,\ddagger} + 2RT\end{aligned}\quad (7.4.53)$$

To summarise,

Theorem 7.4.3.3 (Activation energy in terms of $\Delta_r H^{\circ,\ddagger}$). For a reaction in solution phase,

$$E_a = \Delta_r H^{\circ,\ddagger} + RT \quad (7.4.54)$$

For a bimolecular reaction in gas phase,

$$E_a = \Delta_r H^{\circ,\ddagger} + 2RT \quad (7.4.55)$$

And so we can work out the pre-exponential factor, A , with the help with eq. 7.4.41:

⁷Equation 7.4.45 is technically derived from the concentration equilibrium constant K_c , but as E_a is proportional to $\partial \ln k_{2\text{nd}} / \partial T$, the references to standard states don't make a material difference, but we just need to be aware that we technically need to go from an expression derived from K_p .

Theorem 7.4.3.4 (Pre-exponential factor). We just need to factor out the extra RT term in the expression of $\Delta_r H^{\circ,\ddagger}$ to get, for solution phase

$$A = \frac{kT}{h} \exp \left(\frac{\Delta_r S^{\circ,\ddagger}}{R} + 1 \right) \quad (7.4.56)$$

or in gas phase, for a bimolecular reaction

$$A = \left(\frac{1}{p^\circ} \right) \frac{kT}{h} \exp \left(\frac{\Delta_r S^{\circ,\ddagger}}{R} + 2 \right) \quad (7.4.57)$$

7.4.3.3 Experimental determination of $\Delta_r S^{\circ,\ddagger}$ and $\Delta_r H^{\circ,\ddagger}$

Starting from eq. 7.4.41, we can write

$$\begin{aligned} k_{2\text{nd}} &= \frac{kT}{h} \exp \left(\frac{\Delta_r S^{\circ,\ddagger}}{R} \right) \exp \left(\frac{-\Delta_r H^{\circ,\ddagger}}{RT} \right) \\ \ln \frac{k_{2\text{nd}}}{T} &= \ln \frac{k}{h} + \frac{\Delta_r S^{\circ,\ddagger}}{R} - \frac{\Delta_r H^{\circ,\ddagger}}{R} \left(\frac{1}{T} \right) \end{aligned} \quad (7.4.58)$$

Basically we need to plot $\ln(k_{2\text{nd}}/T)$ against $(1/T)$, the slope is $-\Delta_r H^{\circ,\ddagger}/R$ and the intercept $(\ln k/h + \Delta_r S^{\circ,\ddagger}/R)$. However, usually measurement of $k_{2\text{nd}}$ occurs across a small temperature range, so a long extrapolation is needed to get $\Delta_r S^{\circ,\ddagger}$, so that is likely to be inaccurate.

7.4.3.4 Volume of activation

For a gas-phase reaction, altering the partial pressures of the reactants (essentially concentrations) of the reactants will alter the rate of reaction. However in solution-phase reactions, at high pressures, rate is also affected by the pressure. This can only mean that $k_{2\text{nd}}$ is dependent on p somehow. We remember a definition of volume is

$$V = \left(\frac{\partial G}{\partial p} \right)_T \quad (7.4.59)$$

Starting from eq. 7.4.37 we can write⁸

$$\ln k_{2\text{nd}} = \ln \frac{kT}{h} - \frac{\Delta_r G^{\circ,\ddagger}}{RT}$$

$$\left(\frac{\partial \ln k_{2\text{nd}}}{\partial p} \right)_T = \frac{-1}{RT} \left(\frac{\partial \Delta_r G^{\circ,\ddagger}}{\partial p} \right)_T = \frac{-\Delta_r V^{\circ,\ddagger}}{RT} \quad (7.4.60)$$

As the SI unit of pressure is Pa, $\Delta_r V^{\circ,\ddagger}$ will be in $\text{m}^3 \text{mol}^{-1}$.

Formal integration of the above equation between the standard pressure (10^5 Pa) and an arbitrary pressure, at constant temperature, gives

$$\int_{k_{2\text{nd}}(p^\circ)}^{k_{2\text{nd}}(p)} d(\ln k_{2\text{nd}}) = \frac{-\Delta_r V^{\circ,\ddagger}}{RT} \int_{p^\circ}^p dp$$

$$\ln \left(\frac{k_{2\text{nd}}(p)}{k_{2\text{nd}}(p^\circ)} \right) = \frac{-\Delta_r V^{\circ,\ddagger}}{RT} (p - p^\circ) \quad (7.4.61)$$

Basically, the gradient of a plot of $\ln k_{2\text{nd}}$ against p gives $-\Delta_r V^{\circ,\ddagger}/RT$. $\Delta_r V^{\circ,\ddagger}$ can be either positive or negative and are usually in the range of $10 \text{ cm}^3 \text{mol}^{-1}$.

7.4.3.5 Interpretation of experimental data

With the plots introduced above we can obtain values for the entropy and volume of activation, and depending on whether it's a gas-phase reaction or a solution-phase reaction we can draw some conclusions about the possible reaction mechanism.

Gas-phase reactions

Gas-phase reactions are easier to interpret than solution-phase ones, as solvation effects are not present. Examples include

- **Bimolecular reactions:** $\Delta_r S^{\circ,\ddagger} < 0$ as the translational entropy decreases from two lots to one lot, and a similar but smaller contribution from reduction in rotational entropy.
 - Bimolecular Diels-Alder reaction between butadiene and ethene: $\Delta_r S^{\circ,\ddagger} = -150 \text{ J K}^{-1} \text{mol}^{-1}$, due to massively reduced *translational* entropy.

⁸This is to derive the solution-phase relation, but as we can see, as before, taking derivative makes references to standard state immaterial, so this is applicable to gas-phase reactions as well.

- **Unimolecular reactions:** $\Delta_r S^{\circ,\ddagger} < 0$ as the rotational and vibrational entropies increase. Translational entropy does not change appreciably as the transition state is still a single molecule.
 - Reverse Diels-Alder of dicyclopentadiene: $\Delta_r S^{\circ,\ddagger} = -8 \text{ J K}^{-1} \text{ mol}^{-1}$, no change in translational entropy, due to the concerted nature of Diels-Alder, no additional vibrational and rotational entropies are gained, unlike below.
 - Decomposition of ethane to 2 methyl radicals: $\Delta_r S^{\circ,\ddagger} = +58 \text{ J K}^{-1} \text{ mol}^{-1}$, this is due to increased vibrational and rotational entropies due to the lengthened bond.
- **Intramolecular reactions:** $\Delta_r S^{\circ,\ddagger} < 0$, as internal vibrational and rotational entropies are lost.
 - Intramolecular Diels-Alder: $\Delta_r S^{\circ,\ddagger} = -32 \text{ J K}^{-1} \text{ mol}^{-1}$, due to formation of a constrained ring.

Solution-phase reactions

In non-polar solvents, interpretation usually follows the gas-phase guidelines: bimolecular Diels-Alder have $\Delta_r S^{\circ,\ddagger} = -150 \text{ J K}^{-1} \text{ mol}^{-1}$.

However in polar solvents, we have to consider both the entropy changes of the molecule, *i.e.*, the *intrinsic effects* (IE), and the entropy change of the solvent, *i.e.*, the *solvent effects* (SE). They can agree or oppose each other so careful analysis of data is needed. The effects described above for gas phase reactions are essentially the intrinsic effects. The solvent effect is essentially the extent of solvation, and roughly speaking

$$\Delta_r S^{\circ,\ddagger}(\text{SE}) \propto \frac{q^2}{r} \quad (7.4.62)$$

Therefore the following scenarios can arise

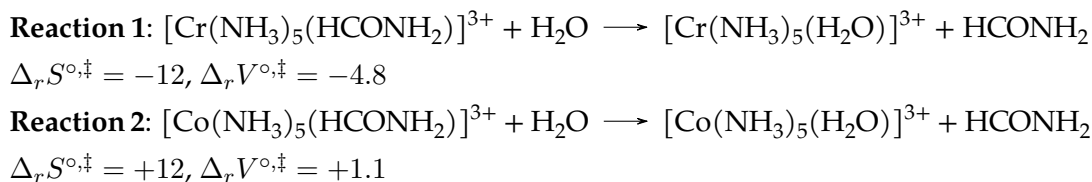
- Charge effects (electrostriction)
 - When two oppositely charged species react, the TS is overall neutral and previously coordinated solvent molecules are released, $\Delta_r S^{\circ,\ddagger} > 0$.
 - When a previously neutral molecule undergoes heterolytic fission and *charge separation* results, essentially creating an electric dipole, therefore the transition state will be more solvated, $\Delta_r S^{\circ,\ddagger} < 0$.

- Ionic radius effects
 - Spreading out of charges: unimolecular decay of benzene diazonium ion giving N_2 and $[\text{C}_6\text{H}_6]^+$, the transition state have the same charge, but more spread out, effectively making the ionic radius larger, which decreases the solvent effect. Another way to look at this is that the electric dipole decreases due to this spreading out of charges. $\Delta_r S^\ddagger = +44 \text{ J K}^{-1} \text{ mol}^{-1}$. The alternative, nucleophilic attack by water would involve a decrease in entropy as it's an associative mechanism.

$\Delta_r S^\ddagger$ and $\Delta_r V^\ddagger$ are well correlated as more or less the same factors determine them. A further set of examples come from ligand exchange in transition metal complexes:

- **Dissociative:** Both volume and entropy of activation positive.
- **Associative:** Both negative
- **Interchange** (concerted): entropy almost zero, but as TS is 'looser', volume could be larger, but in practice could have either sign.

Now we can look at some real examples and how to interpret them, especially on how to untangle the intrinsic and solvent effects.



Analysis

As the entropies of activation are both appreciable, the interchange mechanism can be ruled out, so these are either associative or dissociative.

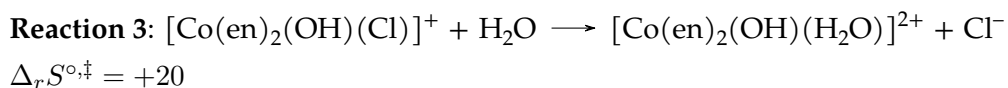
If R1 is associative, the entropy and volume of activation should both be negative *by intrinsic effects*, due to reduced translational entropy. And the solvent effect should be negligible as the dipole do not change significantly upon entering TS

If R1 is dissociative, by intrinsic effects both parameters should be positive, and there's also no appreciable solvent effects as the dipole does not change.

We can thus draw up a table

| | AS | DS |
|----|----|----|
| IE | - | + |
| SE | 0 | 0 |

so only the associative mechanism can explain the data.

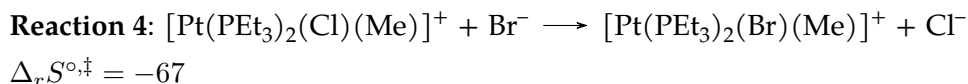


Analysis

We'll draw the same table

| | AS | DS |
|----|----|----|
| IE | - | + |
| SE | 0 | - |

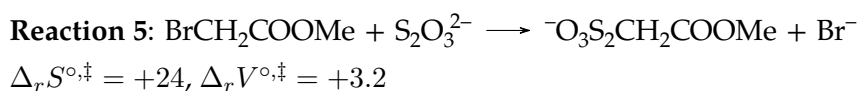
The intrinsic effect for these simple bimolecular reactions is invariable, and only the solvent effect will differ. In this case, of note is the dissociative solvent effects, the dipole is larger upon charge separation, so the TS will be more solvated. The only option that makes sense is a dissociative mechanism where the intrinsic effect predominates the opposing solvent effects.

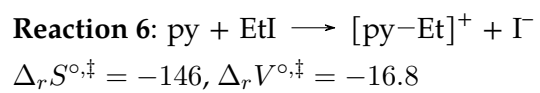


Analysis

| | AS | DS |
|----|----|----|
| IE | - | + |
| SE | + | - |

Now that's a little awkward as in both cases the two effects work against each other. However *intrinsic effect is usually larger than solvent effects*, so we can conclude that associative mechanism is the most likely.





For R5,

| | AS | DS |
|----|----|----|
| IE | - | + |
| SE | 0 | - |

A dissociative mechanism is likely.

For R6,

| | AS | DS |
|----|----|----|
| IE | - | + |
| SE | - | - |

An associative mechanism is likely.

Appendix A

Quick reference

Correspondence between energy and wavenumbers

Setting $\tilde{c}_0 = 100c_0$,

$$1 \text{ J} = \frac{1}{h\tilde{c}_0} = 5.0341 \times 10^{22} \text{ cm}^{-1}$$

| Energy | Wavenumbers |
|------------------|---|
| $\hbar\omega$ | $\tilde{\omega} = \omega/2\pi\tilde{c}_0$ |
| $B = \hbar^2/2I$ | $\tilde{B} = \frac{h}{4\pi I\tilde{c}_0}$ |

Particle in a box solutions (Equation 1.2.33)

$$\psi_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right), \text{ and } E_n = \frac{n^2\pi^2\hbar^2}{2ma^2}$$

Morse potential (Section 1.2.3.3)

Functional form

$$V_M(r) = D_e \left[1 - e^{-\beta(r-r_e)}\right]^2$$

Parameters

$$\beta = \sqrt{\frac{k_e}{2D_e}}$$

$$E_\nu = \left(\nu + \frac{1}{2}\right) \hbar\omega - \left(\nu + \frac{1}{2}\right)^2 \hbar\omega x_e$$

$$\tilde{\epsilon}_\nu = \left(\nu + \frac{1}{2}\right) \tilde{\omega} - \left(\nu + \frac{1}{2}\right)^2 \tilde{\omega} x_e$$

$$D_e = \frac{\hbar\omega}{4x_e}$$

$$\tilde{D}_e = \frac{\tilde{\omega}}{4x_e}$$

$$D_0 = D_e - E_0 = D_e(1 - x_e)^2$$

$$\tilde{D}_0 = \tilde{D}_e(1 - x_e)^2$$

Laplacian in Cartesian and spherical

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \left(\frac{\partial^2}{\partial \phi^2} \right)$$

Angular momentum operators

$$L_z = \frac{\hbar}{i} \frac{\partial}{\partial \phi}$$

$$L^2 = -\hbar^2 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right]$$

Angular momentum commutator

$$[L_x, L_y] = i\hbar L_z$$

Eigenvalues of H , L^2 and L_z

These three operators commute, therefore they have simultaneous eigenfunctions.

$$H\psi = E\psi$$

$$L^2\psi = \hbar l(l+1)\psi$$

$$L_z\psi = \hbar m\psi$$

Functional forms of p and d orbitals

p -orbitals

$$\psi_{2p_x} \equiv \frac{1}{\sqrt{2}}(-\psi_{2p,+1} + \psi_{2p,-1}) \propto xe^{-r/2a}$$

$$\psi_{2p_y} \equiv \frac{1}{\sqrt{2}}i(\psi_{2p,+1} + \psi_{2p,-1}) \propto ye^{-r/2a}$$

$$\psi_{2p_z} \equiv \psi_{2p,0} \propto ze^{-r/2a}$$

d -orbitals

$$\psi_{3d_{z^2}} = \psi_{3d,0} \propto (3z^2 - r^2)e^{-r/3a}$$

$$\psi_{3d_{xz}} = \frac{1}{\sqrt{2}}(\psi_{3d,+1} + \psi_{3d,-1}) \propto xye^{-r/3a}$$

$$\psi_{3d_{yz}} = \frac{1}{\sqrt{2}}i(-\psi_{3d,+1} + \psi_{3d,-1}) \propto yze^{-r/3a}$$

$$\psi_{3d_{x^2-y^2}} = \frac{1}{\sqrt{2}}(\psi_{3d,+2} + \psi_{3d,-2}) \propto (x^2 - y^2)e^{-r/3a}$$

$$\psi_{3d_{xy}} = \frac{1}{\sqrt{2}}i(-\psi_{3d,+2} + \psi_{3d,-2}) \propto xye^{-r/3a}$$

Term symbols

Atomic term symbols

These are in the form of $^{(2S+1)}L_J$

Bibliography

- [1] George B Arfken, Hans J Weber, and Frank E Harris. *Mathematical methods for physicists: a comprehensive guide*. Elsevier, 2012.
- [2] Peter W Atkins and Ronald S Friedman. *Molecular quantum mechanics*. Oxford university press, 2011.
- [3] Barton Zweibach. Chapter 1: Perturbation Theory (PDF) - 8-06 Quantum Mechanics, MIT OpenCourseWare. https://ocw.mit.edu/courses/physics/8-06-quantum-physics-iii-spring-2018/lecture-notes/MIT8_06S18ch1.pdf. [Online; accessed 23-August-2019].
- [4] Barton Zweibach. MIT 8.06 Quantum Physics III, Spring 2018, MIT OpenCourseWare - YouTube. <https://www.youtube.com/playlist?list=PLU14u3cNGP60Zcz8LnCDFI8RPqRhJbb4L>. [Online; accessed 23-August-2019].
- [5] Stephen J Blundell and Katherine M Blundell. *Concepts in thermal physics*. OUP Oxford, 2009.
- [6] Brown, Kevin. The euler-maclaurin formula — mathpages.com. <https://www.mathpages.com/home/kmath643/kmath643.htm>. [Online; accessed 6-August-2019].
- [7] Herbert B Callen. *Thermodynamics and an Introduction to Thermostatistics*. AAPT, 1998.
- [8] Ken Dill and Sarina Bromberg. *Molecular driving forces: statistical thermodynamics in biology, chemistry, physics, and nanoscience*. Garland Science, 2012.
- [9] Emad Noujeim. Is there a derivation/proof of the generating function for hermite polynomials? <https://www.quora.com/Is-there-a-derivation-proof-of-the-generating-function-for-Hermite-polynomials>, 2019. [Online; accessed 16-July-2019].
- [10] DisintegratingByParts (<https://math.stackexchange.com/users/112478/disintegratingbyparts>). Method of solving an associated legendre equation with $m=0$, and m

- not equal to 0 with explanation - mathematics stack exchange. <https://math.stackexchange.com/q/2069983>. Version: 2016-12-29.
- [11] Leon van Dommelen. 5.6 identical particles. http://www.eng.fsu.edu/~dommelen/quantum/style_a/ident.html. [Online; accessed 12-August-2019].
- [12] Leon van Dommelen. D.18 eigenfunctions of commuting operators. http://www.eng.fsu.edu/~dommelen/quantum/style_a/nt_diag.html. [Online; accessed 31-July-2019].
- [13] Donald A McQuarrie, John D Simon, and John Choi. *Physical Chemistry: A Molecular Approach*. Univ Science Books, 1997.
- [14] Leonard L Schiff. *Quantum mechanics, 3rd edition*. McGraw-Hill, 1968.
- [15] Timothy Hughbanks. Antisymmetric wavefunctions: Slater determinants (06/30/16). <https://www.chem.tamu.edu/rgroup/hughbanks/courses/673/handouts/antisymmetry.pdf>. [Online; accessed 17-August-2019].
- [16] John (<https://physics.stackexchange.com/users/99845/john>). Mutual or same set of eigenfunctions if two operators commute - physics stack exchange. <https://physics.stackexchange.com/q/221007>. Version: 2015-11-28.
- [17] Marijn A. M. Versteegh and Dennis Dieks. The Gibbs paradox and the distinguishability of identical particles. *American Journal of Physics*, 79(7):741–746, Jul 2011.
- [18] Wikipedia contributors. Fuchs' theorem — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Fuchs%27_theorem&oldid=820339705, 2018. [Online; accessed 28-July-2019].
- [19] Wikipedia contributors. Associated legendre polynomials — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Associated_Legendre_polynomials&oldid=906480891, 2019. [Online; accessed 30-July-2019].
- [20] Wikipedia contributors. Ehrenfest theorem — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Ehrenfest_theorem&oldid=896346266, 2019. [Online; accessed 26-July-2019].
- [21] Wikipedia contributors. Hermite polynomials — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Hermite_polynomials&oldid=901572627, 2019. [Online; accessed 16-July-2019].
- [22] Wikipedia contributors. Legendre transformation — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Legendre_transformation&oldid=906032488, 2019. [Online; accessed 13-July-2019].

- [23] Wikipedia contributors. Molecular symmetry — Wikipedia, the free encyclopedia. https://en.wikipedia.org/wiki/Molecular_symmetry#Common_point_groups, 2019. [Online; accessed 15-September-2019].
- [24] Wikipedia contributors. Nondimensionalization — Wikipedia, the free encyclopedia. <https://en.wikipedia.org/w/index.php?title=Nondimensionalization&oldid=899171082>, 2019. [Online; accessed 15-July-2019].
- [25] Royce KP Zia, Edward F Redish, and Susan R McKay. Making sense of the legendre transform. *American Journal of Physics*, 77(7):614–622, 2009.