Chapter 4

Magnetism

4.1 Introduction

Magnetism arises from two sources. One is the classical magnetic moment due to a current density $\boldsymbol{j}$:

$$\boldsymbol{m} = \frac{1}{2c}\int d^3r\, \boldsymbol{r}\times\boldsymbol{j}\ . \tag{4.1}$$

The other is the intrinsic spin $\boldsymbol{S}$ of a quantum-mechanical particle (typically the electron):

$$\boldsymbol{m} = g\mu_\circ \boldsymbol{S}/\hbar \qquad ; \qquad \mu_\circ = \frac{q\hbar}{2mc} = \text{magneton}, \tag{4.2}$$

where g is the *g-factor* (duh!). For the electron, $q=-e$ and $\mu_\circ = -\mu_{\rm B}$, where $\mu_{\rm B} = e\hbar/2mc$ is the Bohr magneton.

The Hamiltonian for a single electron is

$$\mathcal{H} = \frac{\boldsymbol{\pi}^2}{2m} + V(\boldsymbol{r}) + \frac{e\hbar}{2mc}\,\boldsymbol{\sigma}\cdot\boldsymbol{H} + \frac{\hbar}{4m^2c^2}\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}V\times\boldsymbol{\pi} + \frac{\hbar^2}{8m^2c^2}\boldsymbol{\nabla}^2V + \frac{(\boldsymbol{\pi}^2)^2}{8m^3c^2} + \ldots\ , \tag{4.3}$$

where $\boldsymbol{\pi} = \boldsymbol{p} + \frac{e}{c}\boldsymbol{A}$. Where did this come from? From the Dirac equation,

$$i\hbar\frac{\partial\Psi}{\partial t} = \begin{pmatrix} mc^2+V & c\,\boldsymbol{\sigma}\cdot\boldsymbol{\pi} \\ c\,\boldsymbol{\sigma}\cdot\boldsymbol{\pi} & -mc^2+V\end{pmatrix}\Psi = E\Psi\ . \tag{4.4}$$

The wavefunction Ψ is a four-component Dirac spinor. Since mc^2 is the largest term for our applications, the upper two components of Ψ are essentially the positive energy components. However, the Dirac Hamiltonian mixes the upper two and lower two components of Ψ. One can 'unmix' them by making a canonical transformation,

$$\mathcal{H} \longrightarrow \mathcal{H}' \equiv e^{iS}\,\mathcal{H}\,e^{-iS}\ , \tag{4.5}$$

where S is Hermitian, to render $\mathcal{H}'$ block diagonal. With $E = mc^2 + \varepsilon$, the effective Hamiltonian is given by (4.3). This is known as the Foldy-Wouthuysen transformation, the

details of which may be found in many standard books on relativistic quantum mechanics and quantum field theory (*e.g.* Bjorken and Drell, Itzykson and Zuber, *etc.*) and are recited in §4.10 below. Note that the Dirac equation leads to $g = 2$. If we go beyond "tree level" and allow for radiative corrections within QED, we obtain a perturbative expansion,

$$g = 2\left\{1 + \frac{\alpha}{2\pi} + \mathcal{O}(\alpha^2)\right\} , \tag{4.6}$$

where $\alpha = e^2/\hbar c \approx 1/137$ is the fine structure constant.[1]

There are two terms in (4.3) which involve the electron's spin:

$$\mathcal{H}_{\rm Z} = \frac{e\hbar}{2mc}\,\boldsymbol{\sigma}\cdot\boldsymbol{H} \qquad \text{(Zeeman term)} \tag{4.7}$$

$$\mathcal{H}_{\rm so} = \frac{\hbar}{4m^2c^2}\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}V\times\left(\boldsymbol{p} + \tfrac{e}{c}\boldsymbol{A}\right) \qquad \text{(spin-orbit interaction)} . \tag{4.8}$$

The numerical value for $\mu_{\rm B}$ is

$$\mu_{\rm B} = \frac{e\hbar}{2mc} = 5.788\times10^{-9}\,{\rm eV/G} \tag{4.9}$$

$$\mu_{\rm B}/k_{\rm B} = 6.717\times10^{-5}\,{\rm K/G} . \tag{4.10}$$

So on the scale of electron volts, laboratory scale fields ($H \lesssim 10^6$ G) are rather small. (And $\sim$2000 times smaller for nucleons!).

The thermodynamic magnetization density is defined through

$$\boldsymbol{M} = -\frac{1}{V}\frac{\partial F}{\partial \boldsymbol{H}} , \tag{4.11}$$

where $F(T, V, \boldsymbol{H}, N)$ is the Helmholtz free energy. The susceptibility is then

$$\chi_{\alpha\beta}(\boldsymbol{r}\,|\,\boldsymbol{r}') = -\frac{1}{V}\,\frac{\partial^2 F}{\partial H^\alpha(\boldsymbol{r})\,\partial H^\beta(\boldsymbol{r}')} \quad . \tag{4.12}$$

When the field $\boldsymbol{H}(\boldsymbol{r}, t)$ is time-dependent, we must use time-dependent perturbation theory to compute the time-dependent susceptibility function,

$$\chi_{\alpha\beta}(\boldsymbol{r}, t\,|\,\boldsymbol{r}', t') = \frac{\delta\big\langle M^\alpha(\boldsymbol{r}, t)\big\rangle}{\delta H^\beta(\boldsymbol{r}', t')} , \tag{4.13}$$

where F is replaced by a suitable generating function in the nonequilibrium case. Note that $\boldsymbol{M}$ has the dimensions of $\boldsymbol{H}$.

[1]Note that with $\mu_{\rm n} = e\hbar/2m_{\rm p}c$ for the nuclear magneton, $g_{\rm p} = 2.793$ and $g_{\rm n} = -1.913$. These results immediately suggest that there is composite structure to the nucleons, *i.e.* quarks.

4.1.1 Absence of Orbital Magnetism within Classical Physics

It is amusing to note that classical statistical mechanics cannot account for orbital magnetism. This is because the partition function is independent of the vector potential, which may be seen by simply shifting the origin of integration for the momentum $\boldsymbol{p}$:

$$Z(\boldsymbol{A}) = \mathsf{Tr}\ e^{-\beta\mathcal{H}} = \int \frac{d^N\boldsymbol{r}\, d^N\boldsymbol{p}}{(2\pi\hbar)^{dN}}\, e^{-\beta\mathcal{H}\left(\{\boldsymbol{p}_i - \frac{q}{c}\boldsymbol{A}(\boldsymbol{r}_i), \boldsymbol{r}_i\}\right)} \tag{4.14}$$

$$= \int \frac{d^N\boldsymbol{r}\, d^N\boldsymbol{p}}{(2\pi\hbar)^{dN}}\, e^{-\beta\mathcal{H}(\{\boldsymbol{p}_i, \boldsymbol{r}_i\})} = Z(\boldsymbol{A}=0)\ . \tag{4.15}$$

Thus, the free energy must be independent of $\boldsymbol{A}$ and hence independent of $\boldsymbol{H} = \boldsymbol{\nabla}\times\boldsymbol{A}$, and $\boldsymbol{M} = -\partial F/\partial\boldsymbol{H} = 0$. This inescapable result is known as the Bohr-von Leeuwen theorem. Of course, classical statistical mechanics can describe magnetism due to intrinsic spin, *e.g.*

$$Z_{\text{Heisenberg}}(\boldsymbol{H}) = \prod_i \int \frac{d\hat{\boldsymbol{\Omega}}_i}{4\pi}\, e^{\beta J \sum_{\langle ij\rangle} \hat{\boldsymbol{\Omega}}_i\cdot\hat{\boldsymbol{\Omega}}_j}\, e^{\beta g\mu_\circ \boldsymbol{H}\cdot\sum_i \hat{\boldsymbol{\Omega}}_i} \tag{4.16}$$

$$Z_{\text{Ising}}(H) = \sum_{\{\sigma_i\}} e^{\beta J \sum_{\langle ij\rangle}\sigma_i\,\sigma_j}\, e^{\beta g\mu_\circ H\sum_i \sigma_i}\ . \tag{4.17}$$

Theories of magnetism generally fall into two broad classes: localized and itinerant. In the localized picture, we imagine a set of individual local moments $\boldsymbol{m}_i$ localized at different points in space (typically, though not exclusively, on lattice sites). In the itinerant picture, we focus on delocalized Bloch states which also carry electron spin.

4.2 Basic Atomic Physics

4.2.1 Single electron Hamiltonian

We start with the single-electron Hamiltonian,

$$\mathcal{H} = \frac{1}{2m}\left(\boldsymbol{p} + \frac{e}{c}\boldsymbol{A}\right)^2 + V(\boldsymbol{r}) + g\mu_{\text{B}}\boldsymbol{H}\cdot\boldsymbol{s}/\hbar + \frac{1}{2m^2c^2}\,\boldsymbol{s}\cdot\boldsymbol{\nabla}V\times\left(\boldsymbol{p} + \tfrac{e}{c}\boldsymbol{A}\right)\ . \tag{4.18}$$

For a single atom or ion in a crystal, let us initially neglect effects due to its neighbors. In that case the potential $V(\boldsymbol{r})$ may be taken to be spherically symmetric, so with $\boldsymbol{l} = \boldsymbol{r}\times\boldsymbol{p}$, the first term in the spin-orbit part of the Hamiltonian becomes

$$\mathcal{H}_{\text{so}} = \frac{1}{2m^2c^2}\,\boldsymbol{s}\cdot\boldsymbol{\nabla}V\times\boldsymbol{p} = \frac{1}{2m^2c^2}\,\frac{1}{r}\frac{\partial V}{\partial r}\boldsymbol{s}\cdot\boldsymbol{l}\ , \tag{4.19}$$

with $\boldsymbol{\nabla}V = \hat{\boldsymbol{r}}(\partial V/\partial r)$. We adopt the gauge $\boldsymbol{A} = \frac{1}{2}\boldsymbol{H}\times\boldsymbol{r}$ so that

$$\frac{1}{2m}\left(\boldsymbol{p} + \frac{e}{c}\boldsymbol{A}\right)^2 = \frac{\boldsymbol{p}^2}{2m} + \frac{e}{2mc}\,\boldsymbol{H}\cdot\boldsymbol{l} + \frac{e^2}{8mc^2}\,(\boldsymbol{H}\times\boldsymbol{r})^2\ . \tag{4.20}$$

Finally, restoring the full SO term, we have

$$\mathcal{H} = \frac{\boldsymbol{p}^2}{2m} + V(\boldsymbol{r}) + \frac{1}{\hbar}\mu_{\rm B}(\boldsymbol{l} + 2\boldsymbol{s}) \cdot \boldsymbol{H} + \frac{1}{2m^2c^2}\,\frac{1}{r}\,\frac{\partial V}{\partial r}\,\boldsymbol{l}\cdot\boldsymbol{s} \tag{4.21}$$

$$+ \frac{e^2}{8mc^2}\,(\boldsymbol{H}\times\boldsymbol{r})^2 + \frac{\mu_{\rm B}}{\hbar}\,\frac{rV'(r)}{4mc^2}\,2\boldsymbol{s}\cdot\big[\boldsymbol{H} - \hat{\boldsymbol{r}}(\boldsymbol{H}\cdot\hat{\boldsymbol{r}})\big]\ . \tag{4.22}$$

The last term is usually negligible because $rV'(r)$ is on the scale of electron volts, while $mc^2 = 511\,\text{keV}$.[2] The $(\boldsymbol{H}\times\boldsymbol{r})^2$ breaks the rotational symmetry of an isolated ion, so in principal we cannot describe states by total angular momentum J. However, this effect is of order $\boldsymbol{H}^2$, so if we only desire energies to order $\boldsymbol{H}^2$, we needn't perturb the wavefunctions themselves with this term, *i.e.* we can simply treat it within first order perturbation theory, leading to an energy shift $\frac{e^2}{8mc^2}\big\langle\,\Psi\,\big|\,(\boldsymbol{H}\times\boldsymbol{r})^2\,\big|\,\Psi\,\big\rangle$ in state $\big|\,n\,\big\rangle$.

4.2.2 The Darwin Term

If $V(\boldsymbol{r}) = -Ze^2/r$, then from $\boldsymbol{\nabla}^2(1/r) = -4\pi\delta(\boldsymbol{r})$ we have

$$\frac{\hbar^2}{8m^2c^2}\,\boldsymbol{\nabla}^2 V = \frac{Z\pi e^2\hbar^2}{2m^2c^2}\,\delta(\boldsymbol{r})\ , \tag{4.23}$$

which is centered at the nucleus. This leads to an energy shift for s-wave states,

$$\Delta E_{s-\text{wave}} = \frac{Z\pi e^2\hbar^2}{2m^2c^2}\,\big|\psi(0)\big|^2 = \frac{\pi}{2}\,Z\,\alpha^2 a_{\rm B}^3\,\big|\psi(0)\big|^2\cdot\frac{e^2}{a_{\rm B}}\ , \tag{4.24}$$

where $\alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137}$ is the fine structure constant and $a_{\rm B} = \frac{\hbar^2}{me^2} \approx 0.529\,\text{Å}$ is the Bohr radius. For large Z atoms and ions, the Darwin term contributes a significant contribution to the total energy.

4.2.3 Many electron Hamiltonian

The full N-electron atomic Hamiltonian, for nuclear charge Ze, is then

$$\mathcal{H} = \sum_{i=1}^N\left[\frac{\boldsymbol{p}_i^2}{2m} - \frac{Ze^2}{r_i}\right] + \sum_{i<j}^N\frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} + \sum_{i=1}^N\zeta(r_i)\,\boldsymbol{l}_i\cdot\boldsymbol{s}_i$$

$$+ \sum_{i=1}^N\left\{\frac{\mu_{\rm B}}{\hbar}\,(\boldsymbol{l}_i + 2\boldsymbol{s}_i)\cdot\boldsymbol{H} + \frac{e^2}{8mc^2}(\boldsymbol{H}\times\boldsymbol{r}_i)^2\right\}\ , \tag{4.25}$$

where $\boldsymbol{l}_i = \boldsymbol{r}_i\times\boldsymbol{p}_i$ and

$$\zeta(r) = \frac{Ze^2}{2m^2c^2}\,\frac{1}{r^3} = \frac{Z}{\hbar^2}\left(\frac{e^2}{\hbar c}\right)^2\frac{e^2}{2a_{\rm B}}\left(\frac{a_{\rm B}}{r}\right)^3\ . \tag{4.26}$$

[2] Exercise: what happens in the case of high Z atoms?

The total orbital and spin angular momentum are $\boldsymbol{L} = \sum_i \boldsymbol{l}_i$ and $\boldsymbol{S} = \sum_i \boldsymbol{s}_i$, respectively.

The full many-electron atom is too difficult a problem to solve exactly. Generally progress is made by using the Hartree-Fock method to reduce the many-body problem to an effective one-body problem. One starts with the interacting Hamiltonian

$$\mathcal{H} = \sum_{i=1}^{N} \left[\frac{\boldsymbol{p}_i^2}{2m} - \frac{Ze^2}{r_i} \right] + \sum_{i<j}^{N} \frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \ , \tag{4.27}$$

and treats $\mathcal{H}_{\rm so}$ as a perturbation, and writes the best possible single Slater determinant state:

$$\Psi_{\sigma_1 \ldots \sigma_N}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = \mathcal{A}\Big[\varphi_{1\sigma_1}(\boldsymbol{r}_1) \cdots \varphi_{N\sigma_N}(\boldsymbol{r}_N) \Big] \ , \tag{4.28}$$

where $\mathcal{A}$ is the antisymmetrizer, and $\varphi_{i\sigma}(\boldsymbol{r})$ is a single particle wavefunction. In second-quantized notation, the Hamiltonian is

$$\mathcal{H} = \sum_{ij\sigma} T_{ij}^{\sigma} \, \psi_{i\sigma}^\dagger \, \psi_{j\sigma} + \sum_{\substack{ijkl \\ \sigma\sigma'}} V_{ijkl}^{\sigma\sigma'} \, \psi_{i\sigma}^\dagger \, \psi_{j\sigma'}^\dagger \, \psi_{k\sigma'} \, \psi_{l\sigma} \ , \tag{4.29}$$

where

$$T_{ij}^{\sigma} = \int d^3\!r \, \varphi_{i\sigma}^*(\boldsymbol{r}) \Big\{ -\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{|\boldsymbol{r}|} \Big\} \varphi_{j\sigma}(\boldsymbol{r}) \tag{4.30}$$

$$V_{ijkl}^{\sigma\sigma'} = \tfrac{1}{2} \int d^3r \int d^3r' \, \varphi_{i\sigma}^*(\boldsymbol{r}) \, \varphi_{j\sigma'}^*(\boldsymbol{r}') \, \frac{e^2}{|\boldsymbol{r} - \boldsymbol{r}'|} \, \varphi_{k\sigma'}(\boldsymbol{r}') \, \varphi_{l\sigma}(\boldsymbol{r}) \ . \tag{4.31}$$

The Hartree-Fock energy is given by a sum over occupied orbitals:

$$E_{\rm HF} = \sum_{i\sigma} T_{ii}^{\sigma} + \sum_{ij\sigma\sigma'} \Big(V_{ijji}^{\sigma\sigma'} - V_{ijij}^{\sigma\sigma'} \, \delta_{\sigma\sigma'} \Big) \ . \tag{4.32}$$

The term $V_{ijji}^{\sigma\sigma'}$ is called the direct Coulomb, or "Hartree" term, and $V_{ijij}^{\sigma\sigma'} \, \delta_{\sigma\sigma'}$ is the exchange term. Introducing Lagrange multipliers $\varepsilon_{i\sigma}$ to enforce normalization of the $\{\varphi_{i\sigma}(\boldsymbol{r})\}$ and subsequently varying with respect to the wavefunctions yields the Hartree-Fock equations:

$$\left. \frac{\delta E_{\rm HF}}{\delta \varphi_{i\sigma}(\boldsymbol{r})} \right|_{\langle\Psi|\Psi\rangle=1} = 0 \qquad \Longrightarrow \tag{4.33}$$

$$\begin{aligned} \varepsilon_{i\sigma} \, \varphi_{i\sigma}(\boldsymbol{r}) = & \Big\{ -\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{r} \Big\} \varphi_{i\sigma}(\boldsymbol{r}) + \sum_{j\neq i,\sigma'}^{\rm OCC} \int d^3r' \, \frac{\big|\varphi_{j\sigma'}(\boldsymbol{r}')\big|^2}{|\boldsymbol{r} - \boldsymbol{r}'|} \, \varphi_{i\sigma}(\boldsymbol{r}) \\ & - \sum_{j\neq i}^{\rm OCC} \int d^3r' \, \frac{\varphi_{j\sigma}^*(\boldsymbol{r}') \, \varphi_{i\sigma}(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \, \varphi_{j\sigma}(\boldsymbol{r}) \ , \end{aligned} \tag{4.34}$$

which is a set of N coupled integro-differential equations. Multiplying by $\varphi_i^*(\boldsymbol{r})$ and integrating, we find

$$\varepsilon_{i\sigma} = T_{ii}^{\sigma} + 2 \sum_{j\sigma'}^{\rm OCC} \Big(V_{ijji}^{\sigma\sigma'} - V_{ijij}^{\sigma\sigma'} \, \delta_{\sigma\sigma'} \Big) \ . \tag{4.35}$$

It is a good approximation to assume that the Hartree-Fock wavefunctions $\varphi_i(\boldsymbol{r})$ are spherically symmetric, *i.e.*

$$\varphi_{i\sigma}(\boldsymbol{r}) = R_{nl}(r)\, Y_{lm}(\theta,\phi)\ , \tag{4.36}$$

independent of σ. We can then classify the single particle states by the quantum numbers $n \in \{1,2,\ldots\}$, $l \in \{0,1,\ldots,n-1\}$, $m_l \in \{-l,\ldots,+l\}$, and $m_s = \pm\frac{1}{2}$. The essential physics introduced by the Hartree-Fock method is that of *screening*. Close to the origin, a given electron senses a potential $-Ze^2/r$ due to the unscreened nucleus. Farther away, though, the nuclear charge is screened by the core electrons, and the potential decays faster than $1/r$. (Within the Thomas-Fermi approximation, the potential at long distances decays as $-\mathcal{C}e^2a_{\rm B}^3/r^4$, where $\mathcal{C} \simeq 100$ is a numerical factor, independent of Z.) Whereas states of different l and identical n are degenerate for the noninteracting hydrogenic atom, when the nuclear potential is screened, states of different l are no longer degenerate. Smaller l means smaller energy, since these states are localized closer to the nucleus, where the potential is large and negative and relatively unscreened. Hence, for a given n, the smaller l states fill up first. For a given l and n there are $(2s+1)\times(2l+1) = 4l+2$ states, labeled by the angular momentum and spin polarization quantum numbers m_l and m_s; this group of orbitals is called a *shell*.

4.2.4 The Periodic Table

Based on the energetics derived from Hartree-Fock[3], we can start to build up the Periodic Table. (Here I follow the pellucid discussion in G. Baym's *Lectures on Quantum Mechanics*, chapter 20.) Start with the lowest energy states, the 1s orbitals. Due to their lower angular momentum and concomitantly lower energy, the 2s states get filled before the 2p states. Filling the 1s, 2s, and 2p shells brings us to Ne, whose configuration is $(1s)^2\,(2s)^2\,(2p)^6$. Next comes the 3s and 3p shells, which hold eight more electrons, and bring us to Ar: $1s^2\,2s^2\,2p^6\,3s^2\,3p^6 = [\mathrm{Ne}]\,3s^2\,3p^6$, where the symbol [Ne] denotes the electronic configuration of neon. At this point, things start to get interesting. The 4s orbitals preempt the 3d orbitals, or at least most of the time. As we see from table 4.1, there are two anomalies

The 3d transition metal series ([Ar] core additions)					
Element (A^Z)	Sc^{21}	Ti^{22}	V^{23}	Cr^{24}	Mn^{25}
Configuration	$4s^2\,3d^1$	$4s^2\,3d^2$	$4s^2\,3d^3$	$4s^1\,3d^5$	$4s^2\,3d^5$
Element (A^Z)	Fe^{26}	Co^{27}	Ni^{28}	Cu^{29}	Zn^{30}
Configuration	$4s^2\,3d^6$	$4s^2\,3d^7$	$4s^2\,3d^8$	$4s^1\,3d^{10}$	$4s^2\,3d^{10}$

Table 4.1: Electronic configuration of 3d-series metals.

in the otherwise orderly filling of the 3d shell. Chromium's configuration is $[\mathrm{Ar}]\,4s^1\,3d^5$ rather than the expected $[\mathrm{Ar}]\,4s^2\,3d^4$, and copper's is $[\mathrm{Ar}]\,4s^1\,3d^{10}$ and not $[\mathrm{Ar}]\,4s^2\,3d^9$. In

[3]Hartree-Fock theory tends to overestimate ground state atomic energies by on the order of 1 eV per pair of electrons. The reason is that electron-electron correlations are not adequately represented in the Hartree-Fock many-body wavefunctions, which are single Slater determinants.

reality, the ground state is not a single Slater determinant and involves linear combinations of different configurations. But the largest weights are for Cr and Cu configurations with only one 4s electron. Zinc terminates the 3d series, after which we get orderly filling of the 4p orbitals.

Row five reiterates row four, with the filling of the 5s, 4d, and 5p shells. In row six, the lanthanide (4f) series interpolates between 6s and 5d (see also table 4.2), and the actinide (5f) series interpolates in row seven between 7s and 6d.

Shell:	1s	2s	2p	3s	3p	4s	3d	4p	5s
Termination:	^{2}He	^{4}Be	^{10}Ne	^{12}Mg	^{18}Ar	^{20}Ca	^{30}Zn	^{36}Kr	^{38}Sr
Shell:	4d	5p	6s	4f	5d	6p	7s	5f/6d	
Termination:	^{48}Cd	^{54}Xe	^{56}Ba	^{71}Lu	^{80}Hg	^{86}Rn	^{88}Ra	^{102}No	

Table 4.2: Rough order in which shells of the Periodic Table are filled.

4.2.5 Splitting of Configurations: Hund's Rules

The electronic configuration does not uniquely specify a ground state. Consider, for example, carbon, whose configuration is $1s^2\,2s^2\,2p^2$. The filled 1s and 2s shells are inert. However, there are $\binom{6}{2} = 15$ possible ways to put two electrons in the 2p shell. It is convenient to label these states by total L, S, and J quantum numbers, where $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$ is the total angular momentum. It is standard to abbreviate each such multiplet with the label $^{2S+1}L_J$, where L = S, P, D, F, H, *etc.*. For carbon, the largest L value we can get is $L = 2$, which requires $S = 0$ and hence $J = L = 2$. This 5-fold degenerate multiplet is then abbreviated $^1\mathrm{D}_2$. But we can also add together two $l = 1$ states to get total angular momentum $L = 1$ as well. The corresponding spatial wavefunction is antisymmetric, hence $S = 1$ in order to achieve a symmetric spin wavefunction. Since $|L - S| \le J \le |L + S|$ we have $J = 0$, $J = 1$, or $J = 2$ corresponding to multiplets $^3\mathrm{P}_0$, $^3\mathrm{P}_1$, and $^3\mathrm{P}_2$, with degeneracy 1, 3, and 5, respectively. The final state has $J = L = S = 0$: $^1\mathrm{S}_0$. The Hilbert space is then spanned by two $J = 0$ singlets, one $J = 1$ triplet, and two $J = 2$ quintuplets: $0 \oplus 0 \oplus 1 \oplus 2 \oplus 2$. That makes 15 states. Which of these is the ground state?

The ordering of the multiplets is determined by the famous *Hund's rules*:

1. The LS multiplet with the *largest* S has the lowest energy.

2. If the largest value of S is associated with several multiplets, the multiplet with the *largest* L has the lowest energy.

3. If an incomplete shell is not more than half-filled, then the lowest energy state has $J = |L - S|$. If the shell is more than half-filled, then $J = L + S$.

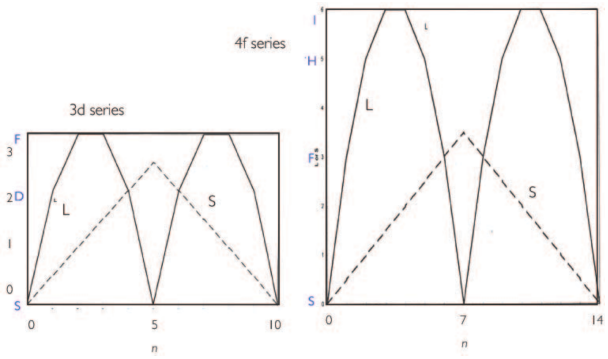

Figure 4.1: Variation of L, S, and J among the 3d and 4f series.

Hund's rules are largely empirical, but are supported by detailed atomic quantum many-body calculations. Basically, rule #1 prefers large S because this makes the spin part of the wavefunction maximally symmetric, which means that the spatial part is maximally anti-symmetric. Electrons, which repel each other, prefer to exist in a spatially antisymmetric state. As for rule #2, large L expands the electron cloud somewhat, which also keeps the electrons away from each other. For neutral carbon, the ground state has $S = 1$, $L = 1$, and $J = |L - S| = 0$, hence the ground state term is ${}^3\mathrm{P}_0$.

Let's practice Hund's rules on a couple of ions:

- P: The electronic configuration for elemental phosphorus is $[\mathrm{Ne}]\,3\mathrm{s}^2\,3\mathrm{p}^3$. The unfilled 3d shell has three electrons. First maximize S by polarizing all spins parallel (up, say), yielding $S = \frac{3}{2}$. Next maximize L consistent with Pauli exclusion, which says $L = -1 + 0 + 1 = 0$. Finally, since the shell is exactly half-filled, and not more, $J = |L - S| = \frac{3}{2}$, and the ground state term is ${}^4\mathrm{S}_{3/2}$.

- Mn^{4+}: The electronic configuration $[\mathrm{Ar}]\,4\mathrm{s}^0\,3\mathrm{d}^3$ has an unfilled 3d shell with three electrons. First maximize S by polarizing all spins parallel, yielding $S = \frac{3}{2}$. Next maximize L consistent with Pauli exclusion, which says $L = 2 + 1 + 0 = 3$. Finally, since the shell is less than half-filled, $J = |L - S| = \frac{3}{2}$. The ground state term is ${}^4\mathrm{F}_{3/2}$.

- Fe^{2+}: The electronic configuration $[\mathrm{Ar}]\,4\mathrm{s}^0\,3\mathrm{d}^6$ has an unfilled 3d shell with six electrons, or four holes. First maximize S by making the spins of the holes parallel, yielding $S = 2$. Next, maximize L consistent with Pauli exclusion, which says $L = 2 + 1 + 0 + (-1) = 2$ (adding L_z for the four holes). Finally, the shell is more than half-filled, which means $J = L + S = 4$. The ground state term is ${}^5\mathrm{D}_4$.

- Nd^{3+}: The electronic configuration $[\mathrm{Xe}]\,6\mathrm{s}^0\,4\mathrm{f}^3$ has an unfilled 4f shell with three electrons. First maximize S by making the electron spins parallel, yielding $S = \frac{3}{2}$. Next, maximize L consistent with Pauli exclusion: $L = 3 + 2 + 1 = 6$. Finally, the shell is less than half-filled, we have $J = |L - S| = \frac{9}{2}$. The ground state term is ${}^4\mathrm{I}_{9/2}$.

4.2.6 Spin-Orbit Interaction

Hund's third rule derives from an analysis of the spin-orbit Hamiltonian,

$$\mathcal{H}_{\rm so} = \sum_{i=1}^{N} \zeta(r_i)\, \boldsymbol{l}_i \cdot \boldsymbol{s}_i \ . \tag{4.37}$$

This commutes with $\boldsymbol{J}^2$, $\boldsymbol{L}^2$, and $\boldsymbol{S}^2$, so we can still classify eigenstates according to total J, L, and S. The Wigner-Eckart theorem then guarantees that within a given J multiplet, we can replace any tensor operator transforming as

$$\mathcal{R}\, T_{JM}\, \mathcal{R}^\dagger = \sum_{M'} \mathcal{D}^J_{MM'}(\alpha,\beta,\gamma)\, T_{JM'} \ , \tag{4.38}$$

where $\mathcal{R}$ corresponds to a rotation through Euler angles α, β, and γ, by a product of a reduced matrix element and a Clebsch-Gordon coefficient:

$$\big\langle\, JM \,\big|\, T_{\boldsymbol{J}''M''} \,\big|\, J'M' \,\big\rangle = C\begin{pmatrix} J & J' & J'' \\ M & M' & M'' \end{pmatrix} \big\langle J \big|\big| T_{J''} \big|\big| J' \big\rangle \ . \tag{4.39}$$

In other words, if two tensor operators have the same rank, their matrix elements are proportional. Both $\mathcal{H}_{\rm so}$ and $\boldsymbol{L}\cdot\boldsymbol{S}$ are products of rank $L=1$, $S=1$ tensor operators, hence we may replace

$$\mathcal{H}_{\rm so} \longrightarrow \tilde{\mathcal{H}}_{\rm so} = \Lambda\, \boldsymbol{L}\cdot\boldsymbol{S} \ , \tag{4.40}$$

where $\Lambda = \Lambda(N,L,S)$ must be computed from, say, the expectation value of $\mathcal{H}_{\rm so}$ in the state $\big|\, JLSJ \,\big\rangle$. This requires detailed knowledge of the atomic many-body wavefunctions. However, once Λ is known, the multiplet splittings are easily obtained:

$$\begin{aligned} \tilde{\mathcal{H}}_{\rm so} &= \tfrac{1}{2}\Lambda\big(\boldsymbol{J}^2 - \boldsymbol{L}^2 - \boldsymbol{S}^2\big) \\ &= \tfrac{1}{2}\hbar^2\,\Lambda\,\big(J(J+1) - L(L+1) - S(S+1)\big) \ . \end{aligned} \tag{4.41}$$

Thus,

$$E(N,L,S,J) - E(N,L,S,J-1) = \Lambda\, J\, \hbar^2 \ . \tag{4.42}$$

If we replace $\zeta(r_i)$ by its average, then we can find Λ by the following argument. If the last shell is not more than half filled, then by Hund's first rule, the spins are all parallel. Thus $S = \frac{1}{2}N$ and $\boldsymbol{s}_i = \boldsymbol{S}/N$, whence $\Lambda = \langle\zeta\rangle/2S$. Finding $\langle\zeta\rangle$ is somewhat tricky. For $Z^{-1} \ll r/a_{\rm B} \ll 1$, one can use the WKB method to obtain $\psi(r = a_{\rm B}/Z) \sim \sqrt{Z}$, whence

$$\langle\zeta\rangle \sim \left(\frac{Ze^2}{\hbar c}\right)^2 \frac{me^4}{\hbar^4} \tag{4.43}$$

and

$$\Lambda \sim Z^2\,\alpha^2\,\hbar^{-2}\,{\rm Ry} \ , \tag{4.44}$$

where $\alpha = e^2/\hbar c \simeq 1/137$. For heavy atoms, $Z\alpha \sim 1$ and the energy is on the order of that for the outer electrons in the atom.

For shells which are more than half filled, we treat the problem in terms of the *holes* relative to the filled shell case. Since filled shells are inert,

$$\mathcal{H}_{\rm so} = -\sum_{j=1}^{N_{\rm h}} \tilde{\boldsymbol{l}}_j \cdot \tilde{\boldsymbol{s}}_j \ , \tag{4.45}$$

where $N_{\rm h} = 4l + 2 - N$. $\tilde{\boldsymbol{l}}_j$ and $\tilde{\boldsymbol{s}}_j$ are the orbital and spin angular momenta of the holes; $\boldsymbol{L} = -\sum_j \tilde{\boldsymbol{l}}_j$ and $\boldsymbol{S} = -\sum_j \tilde{\boldsymbol{s}}_j$. We then conclude $\Lambda = -\langle\zeta\rangle/2S$. Thus, we arrive at Hund's third rule, which says

$$N \le 2L+1 \quad (\le \text{half-filled}) \qquad \Rightarrow \quad \Lambda > 0 \quad \Rightarrow \quad J = |L-S| \tag{4.46}$$

$$N > 2L+1 \quad (> \text{half-filled}) \qquad \Rightarrow \quad \Lambda < 0 \quad \Rightarrow \quad J = |L+S| \ . \tag{4.47}$$

4.2.7 Crystal Field Splittings

Consider an ion with a single d electron (*e.g.* Cr^{3+}) or a single d hole (*e.g.* Cu^{2+}) in a cubic or octahedral environment. The 5-fold degeneracy of the d levels is lifted by the crystal electric field. Suppose the atomic environment is octahedral, with anions at the vertices of the octahedron (typically O^{2-} ions). In order to minimize the Coulomb repulsion between the d electron and the neighboring anions, the $d_{x^2-y^2}$ and $d_{3x^2-r^2}$ orbitals are energetically disfavored, and this doublet lies at higher energy than the $\{d_{xy}, d_{xz}, d_{yz}\}$ triplet.

The crystal field potential is crudely estimated as

$$V_{\rm CF} = \sum_{\boldsymbol{R}}^{\rm (nbrs)} V(\boldsymbol{r} - \boldsymbol{R}) \ , \tag{4.48}$$

where the sum is over neighboring ions, and V is the atomic potential.

The angular dependence of the cubic crystal field states may be written as follows:

$$\begin{aligned}
d_{x^2-y^2}(\hat{\boldsymbol{r}}) &= \tfrac{1}{\sqrt{2}} Y_{2,2}(\hat{\boldsymbol{r}}) + \tfrac{1}{\sqrt{2}} Y_{2,-2}(\hat{\boldsymbol{r}}) \\
d_{3z^2-r^2}(\hat{\boldsymbol{r}}) &= Y_{2,0}(\hat{\boldsymbol{r}}) \\
d_{xy}(\hat{\boldsymbol{r}}) &= \tfrac{i}{\sqrt{2}} Y_{2,-2}(\hat{\boldsymbol{r}}) - \tfrac{i}{\sqrt{2}} Y_{2,2}(\hat{\boldsymbol{r}}) \\
d_{xz}(\hat{\boldsymbol{r}}) &= \tfrac{1}{\sqrt{2}} Y_{2,1}(\hat{\boldsymbol{r}}) + \tfrac{1}{\sqrt{2}} Y_{2,-1}(\hat{\boldsymbol{r}}) \\
d_{yz}(\hat{\boldsymbol{r}}) &= \tfrac{i}{\sqrt{2}} Y_{2,-1}(\hat{\boldsymbol{r}}) - \tfrac{i}{\sqrt{2}} Y_{2,1}(\hat{\boldsymbol{r}}) \ .
\end{aligned} \tag{4.49}$$

Note that all of these wavefunctions are *real*. This means that the expectation value of L^z, and hence of general L^α, must vanish in any of these states. This is related to the phenomenon of *orbital quenching*, discussed below.

If the internal Hund's rule exchange energy $J_{\rm H}$ which enforces maximizing S is large compared with the ground state crystal field splitting Δ, then Hund's first rule is unaffected.

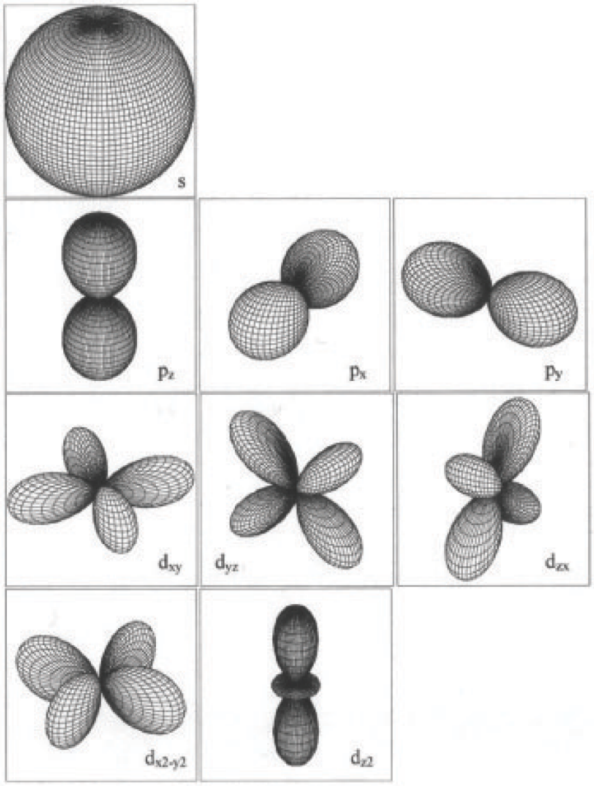

Figure 4.2: Effect on s, p, and d levels of a cubic crystal field.

However, there are examples of ions such as Co^{4+} for which $J_{\rm H} < V_{\rm CF}$. In such cases, the crystal field splitting wins and the ionic ground state is a *low spin state*. For Co^{4+} in an octahedral crystal field, the five 3d electrons all pile into the lower 3-fold degenerate t_{2g} manifold, and the spin is $S = \frac{1}{2}$. When the Hund's rule energy wins, the electrons all have parallel spin and $S = \frac{5}{2}$, which is the usual *high spin state*.

4.3 Magnetic Susceptibility of Atomic and Ionic Systems

To compute the susceptibility, we will need to know magnetic energies to order $\boldsymbol{H}^2$. This can be computed via perturbation theory. Treating the $\boldsymbol{H} = 0$ Hamiltonian as $\mathcal{H}_0$, we have

$$E_n(\boldsymbol{H}) = E_n(0) + \frac{1}{\hbar}\mu_{\rm B}\boldsymbol{H}\cdot\left\langle\, n \,\middle|\, \boldsymbol{L} + 2\boldsymbol{S} \,\middle|\, n \,\right\rangle + \frac{e^2}{8mc^2}\left\langle\, n \,\middle|\, \sum_{i=1}^{Z_{\rm ion}} (\boldsymbol{H}\times\boldsymbol{r}_i)^2 \,\middle|\, n \,\right\rangle$$
$$+ \frac{1}{\hbar^2}\,\mu_{\rm B}^2\, H^\alpha\, H^\beta \sum_{n'\neq n} \frac{\left\langle\, n \,\middle|\, L^\alpha + 2S^\alpha \,\middle|\, n' \,\right\rangle \left\langle\, n' \,\middle|\, L^\beta + 2S^\beta \,\middle|\, n \,\right\rangle}{E_n - E_{n'}} + \mathcal{O}(H^3)\ , \quad (4.50)$$

where $Z_{\rm ion}$ is the number of electrons on the ion or atom in question. Since the $(\boldsymbol{H}\times\boldsymbol{r}_i)^2$ Larmor term is already second order in the field, its contribution can be evaluated in

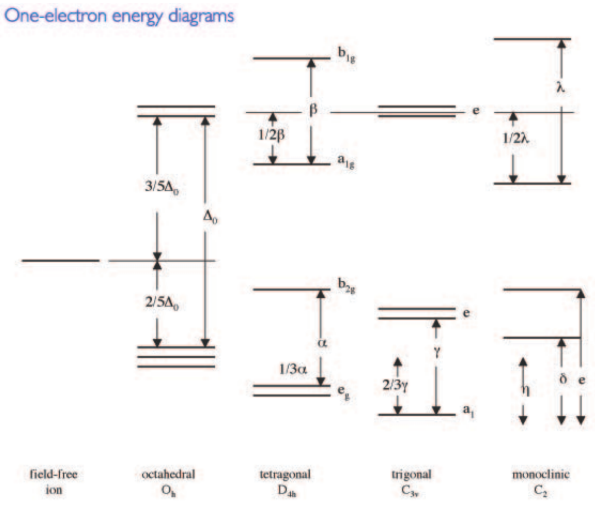

Figure 4.3: The splitting of one-electron states in different crystal field environments.

first order perturbation theory, *i.e.* by taking its expectation value in the state $\big|\, n \,\big\rangle$. The $(\boldsymbol{L}+2\boldsymbol{S})\cdot\boldsymbol{H}$ term, which is linear in the field, is treated in second order perturbation theory.

4.3.1 Filled Shells: Larmor Diamagnetism

If the ground state $\big|\, \mathrm{G} \,\big\rangle$ is a singlet with $\boldsymbol{J}\big|\, \mathrm{G} \,\big\rangle = \boldsymbol{L}\big|\, \mathrm{G} \,\big\rangle = \boldsymbol{S}\big|\, \mathrm{G} \,\big\rangle = 0$, corresponding to a filled shell configuration, then the only contribution to the ground state energy shift is from the Larmor term,

$$\Delta E_0(\boldsymbol{H}) = \frac{e^2\boldsymbol{H}^2}{12mc^2} \left\langle\, \mathrm{G} \,\right|\, \sum_{i=1}^{Z_{\rm ion}} \boldsymbol{r}_i^2 \,\left|\, \mathrm{G} \,\right\rangle \,, \tag{4.51}$$

and the susceptibility is

$$\chi = -\frac{N}{V}\,\frac{\partial^2 \Delta E_0}{\partial H^2} = -\frac{ne^2}{6mc^2} \left\langle\, \mathrm{G} \,\right|\, \sum_{i=1}^{Z_{\rm ion}} \boldsymbol{r}_i^2 \,\left|\, \mathrm{G} \,\right\rangle \,, \tag{4.52}$$

where $n = N/V$ is the density of ions or atoms in question. The sum is over all the electrons in the ion or atom. Defining the mean square ionic radius as

$$\langle \boldsymbol{r}^2 \rangle \equiv \frac{1}{Z_{\rm ion}} \left\langle\, \mathrm{G} \,\right|\, \sum_{i=1}^{Z_{\rm ion}} \boldsymbol{r}_i^2 \,\left|\, \mathrm{G} \,\right\rangle \,, \tag{4.53}$$

we obtain

$$\chi = -\frac{ne^2}{6mc^2} Z_{\rm ion} \langle \boldsymbol{r}^2 \rangle = -\tfrac{1}{6} Z_{\rm ion}\, n a_{\rm B}^3 \left(\frac{e^2}{\hbar c}\right)^2 \frac{\langle r^2 \rangle}{a_{\rm B}^2} \,. \tag{4.54}$$

Note that χ is dimensionless. One defines the *molar susceptibility* as

$$\chi^{\rm molar} \equiv N_{\rm A}\chi/n = -\tfrac{1}{6}Z_{\rm ion}N_{\rm A}a_{\rm B}^3\left(\frac{e^2}{\hbar c}\right)^2\left\langle(\boldsymbol{r}/a_{\rm B})^2\right\rangle$$
$$= -7.91\times10^{-7}\,Z_{\rm ion}\left\langle(\boldsymbol{r}/a_{\rm B})^2\right\rangle{\rm cm}^3/{\rm mol}\ . \tag{4.55}$$

Typically, $\langle(\boldsymbol{r}/a_{\rm B})^2\rangle \sim 1$. Note that with $na_{\rm B}^3 \simeq 0.1$, we have $|\chi| \lesssim 10^{-5}$ and $M = \chi H$ is much smaller than H itself.

Molar Susceptibilities of Noble Gas Atoms and Alkali and Halide Ions					
Atom or Ion	Molar Susceptibility	Atom or Atom or Ion	Molar Susceptibility	Atom or Atom or Ion	Molar Susceptibility
		He	-1.9	Li^+	-0.7
F^-	-9.4	Ne	-7.2	Na^+	-6.1
Cl^-	-24.2	Ar	-19.4	K^+	-14.6
Br^-	-34.5	Kr	-28	Rb^+	-22.0
I^-	-50.6	Xe	-43	Cs^+	-35.1

Table 4.3: Molar susceptibilities, in units of 10^{-6} cm^3/mol, of noble gas atoms and alkali and halide ions. (See R. Kubo and R. Nagamiya, eds., *Solid State Physics*, McGrow-Hill, 1969, p. 439.)

4.3.2 Partially Filled Shells: van Vleck Paramagnetism

There are two cases to consider here. The first is when $J = 0$, which occurs whenever the last shell is one electron short of being half-fille. Examples include Eu^{3+} ($4f^6$), Cr^{2+} ($3d^4$), Mn^{3+} ($3d^4$), *etc.* In this case, the first order term vanishes in ΔE_0, and we have

$$\chi = -\frac{ne^2}{6mc^2}\left\langle\,{\rm G}\,\right|\,\sum_{i=1}^{Z_{\rm ion}}\boldsymbol{r}_i^2\,\left|\,{\rm G}\,\right\rangle + 2\,n\,\mu_{\rm B}^2\sum_{n\neq0}\frac{\left|\left\langle\,n\,\right|\,L^z+2S^z\,\left|\,{\rm G}\,\right\rangle\right|^2}{E_n-E_0}\ . \tag{4.56}$$

The second term is positive, favoring alignment of $\boldsymbol{M}$ with $\boldsymbol{H}$. This is called *van Vleck paramagnetism*, and competes with the Larmor diamagnetism.

The second possibility is $J > 0$, which occurs in all cases except filled shells and shells which are one electron short of being half-filled. In this case, the first order term is usually dominant. We label the states by the eigenvalues of the commuting observables $\{\boldsymbol{J}^2, J^z, \boldsymbol{L}^2, \boldsymbol{S}^2\}$. From the Wigner-Eckart theorem, we know that

$$\left\langle\,JLSJ_z\,\right|\,\boldsymbol{L}+2\boldsymbol{S}\,\left|\,JLSJ'_z\,\right\rangle = g_{\rm L}(J,L,S)\left\langle\,JLSJ_z\,\right|\,\boldsymbol{J}\,\left|\,JLSJ'_z\,\right\rangle\ , \tag{4.57}$$

where

$$g_{\rm L}(J,L,S) = \tfrac{3}{2} + \frac{S(S+1)-L(L+1)}{2J(J+1)} \tag{4.58}$$

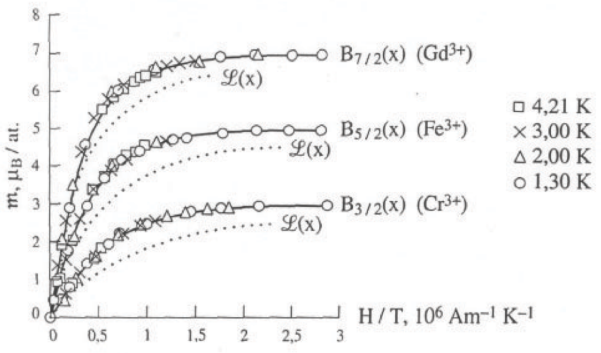

Figure 4.4: Reduced magnetization curves for three paramagnetic salts and comparison with Brillouin theory predictions. $\mathcal{L}(x) = B_{J\to\infty}(x) = \operatorname{ctnh}(x) - x^{-1}$ is the Langevin function.

is known as the Landé g-factor. Thus, the effective Hamiltonian is

$$\mathcal{H}_{\rm eff} = g_{\rm L}\,\mu_{\rm B}\,\boldsymbol{J}\cdot\boldsymbol{H}/\hbar\ . \tag{4.59}$$

The eigenvalues of $\mathcal{H}_{\rm eff}$ are $E_j = j\,\gamma\,H$, where $j \in \{-J,\ldots,+J\}$ and $\gamma = g_{\rm L}\,\mu_{\rm B}$. The problem is reduced to an elementary one in statistical mechanics. The partition function is

$$Z = e^{-F/k_{\rm B}T} = \sum_{j=-J}^{J} e^{-j\gamma H/k_{\rm B}T} = \frac{\sinh\left((J+\frac{1}{2})\gamma H/k_{\rm B}T\right)}{\sinh\left(\gamma H/2k_{\rm B}T\right)}\ . \tag{4.60}$$

The magnetization density is

$$M = -\frac{N}{V}\frac{\partial F}{\partial H} = n\gamma J\,B_J(J\gamma H/k_{\rm B}T)\ , \tag{4.61}$$

where $B_J(x)$ is the *Brillouin function*,

$$B_J(x) = \left(1+\tfrac{1}{2J}\right)\operatorname{ctnh}\left[\left(1+\tfrac{1}{2J}\right)x\right] - \tfrac{1}{2J}\operatorname{ctnh}\left(x/2J\right)\ . \tag{4.62}$$

The magnetic susceptibility is thus

$$\begin{aligned}\chi(H,T) &= \frac{\partial M}{\partial H} = \frac{nJ^2\gamma^2}{k_{\rm B}T}\,B_J'(J\gamma H/k_{\rm B}T) \\ &= (Jg_{\rm L})^2\,(na_{\rm B}^3)\,(e^2/\hbar c)^2\left(\frac{e^2/a_{\rm B}}{k_{\rm B}T}\right)B_J'(g\mu_{\rm B}JH/k_{\rm B}T)\end{aligned} \tag{4.63}$$

$$\chi(H=0,T) = \tfrac{1}{3}(g_{\rm L}\mu_{\rm B})^2\,n\,\frac{J(J+1)}{k_{\rm B}T}\ . \tag{4.64}$$

The inverse temperature dependence is known as *Curie's law*.

Does Curie's law work in solids? The $1/T$ dependence is very accurately reflected in insulating crystals containing transition metal and rare earth ions. We can fit the coefficient of the $1/T$ behavior by defining the 'magneton number' p according to

$$\chi(T) = n\mu_{\rm B}^2\,\frac{p^2}{3k_{\rm B}T}\ . \tag{4.65}$$

The theory above predicts

$$p = g_{\rm L}\sqrt{J(J+1)}\ . \tag{4.66}$$

One finds that the theory works well in the case of rare earth ions in solids. There, the 4f electrons of the rare earths are localized in the vicinity of the nucleus, and do not hybridize significantly with orbitals from neighboring ions. In transition metal compounds, however,

Calculated and Measured Magneton Numbers of Rare Earth Ions				
Ion	Electronic Configuration	Ground State Term $^{(2S+1)}L_J$	magneton $p_{\rm theory}$	magneton $p_{\rm expt}$
La^{3+}	[Xe] $4f^0$	$^1S_0$	0.00	< 0
Ce^{3+}	[Xe] $4f^1$	$^2F_{5/2}$	2.54	2.4
Pr^{3+}	[Xe] $4f^2$	$^3H_4$	3.58	3.5
Nd^{3+}	[Xe] $4f^3$	$^4I_{9/2}$	3.62	3.5
Pm^{3+}	[Xe] $4f^4$	$^5I_4$	2.68	–
Sm^{3+}	[Xe] $4f^5$	$^6H_{5/2}$	0.84	1.5
Eu^{3+}	[Xe] $4f^6$	$^7F_0$	0.00	3.4
Gd^{3+}	[Xe] $4f^7$	$^8S_{7/2}$	7.94	8.0
Tb^{3+}	[Xe] $4f^8$	$^7F_6$	9.72	9.5
Dy^{3+}	[Xe] $4f^9$	$^6H_{15/2}$	10.63	10.6
Ho^{3+}	[Xe] $4f^{10}$	$^5I_8$	10.60	10.4
Er^{3+}	[Xe] $4f^{11}$	$^4I_{15/2}$	9.59	9.5
Tm^{3+}	[Xe] $4f^{12}$	$^3H_6$	7.57	7.3
Yb^{3+}	[Xe] $4f^{13}$	$^2F_{7/2}$	4.54	4.5
Lu^{3+}	[Xe] $4f^{14}$	$^1S_0$	0.00	< 0

Table 4.4: Calculated and measured effective magneton numbers p for rare earth ions. (From N. W. Ashcroft and N. D. Mermin, *Solid State Physics*.) The discrepancy in the cases of Sm and Eu is due to the existence of low-lying multiplets above the ground state.

one finds poor agreement except in the case of S states ($L = 0$). This is because crystal field effects quench the orbital angular momentum, effectively rendering $L = 0$. Indeed, as shown in Table 4.5, the theory can be rescued if one ignores the ground state terms obtained by Hund's rules, and instead takes $L = 0$ and $J = S$, yielding $g_{\rm L} = 2$.

4.4 Itinerant Magnetism of Noninteracting Systems

4.4.1 Pauli Paramagnetism

In a metal, the conduction electrons are delocalized and described by Block states. If we ignore the orbital effects of the magnetic field, we can easily compute the susceptibility at low fields and temperatures. At $T = 0$ and $H = 0$, $\uparrow$ and $\downarrow$ electrons fill respective Fermi

Calculated and Measured Magneton Numbers of Transition Metal Ions					
Ion	Electronic Configuration	Ground State Term $^{(2S+1)}L_J$	magneton $p^{J=\|L\pm S\|}_{\rm theory}$	magneton $p^{J=S}_{\rm theory}$	magneton $p_{\rm expt}$
Ti^{3+}	[Ar] $3d^1$	$^2D_{3/2}$	1.55	1.73	–
V^{4+}	[Ar] $3d^1$	$^2D_{3/2}$	1.55	1.73	1.8
V^{3+}	[Ar] $3d^2$	$^3F_2$	1.63	2.83	2.8
V^{2+}	[Ar] $3d^3$	$^4F_{3/2}$	0.77	3.87	3.8
Cr^{3+}	[Ar] $3d^3$	$^4F_{3/2}$	0.77	3.87	3.7
Mn^{4+}	[Ar] $3d^3$	$^4F_{3/2}$	0.77	3.87	4.0
Cr^{2+}	[Ar] $3d^4$	$^5D_0$	0.00	4.90	4.8
Mn^{3+}	[Ar] $3d^4$	$^5D_0$	0.00	4.90	5.0
Mn^{2+}	[Ar] $3d^5$	$^6S_{5/2}$	5.92	5.92	5.9
Fe^{3+}	[Ar] $3d^5$	$^6S_{5/2}$	5.92	5.92	5.9
Fe^{2+}	[Ar] $3d^6$	$^5D_4$	6.70	4.90	5.4
Co^{2+}	[Ar] $3d^7$	$^4F_{9/2}$	6.54	3.87	4.8
Ni^{2+}	[Ar] $3d^8$	$^3F_4$	5.59	2.83	3.2
Cu^{2+}	[Ar] $3d^9$	$^2D_{5/2}$	3.55	1.73	1.9

Table 4.5: Calculated and measured effective magneton numbers p for transition metal ions. (From N. W. Ashcroft and N. D. Mermin, *Solid State Physics*.) Due to the orbital quenching, the angular momentum is effectively $L = 0$.

seas out to wavevector $k_{\rm F}$. In an external field $\boldsymbol{H}$, the Zeeman interaction splits the energies of the different polarization states:

$$\mathcal{H}_{\rm Z} = \mu_{\rm B}\boldsymbol{\sigma}\cdot\boldsymbol{H}\ . \tag{4.67}$$

Taking $\boldsymbol{H} = H\hat{\boldsymbol{z}}$, and summing over all electrons, the Zeeman Hamiltonian becomes

$$\mathcal{H}_{\rm Z} = \mu_{\rm B}\,H\,(N_\uparrow - N_\downarrow)\ , \tag{4.68}$$

and the magnetization density is (still at $T = 0$)

$$M = -\frac{1}{V}\frac{\partial\mathcal{H}_{\rm Z}}{\partial H} = \mu_{\rm B}\,\frac{N_\downarrow - N_\uparrow}{V}\ . \tag{4.69}$$

Now since the energies of the $\uparrow$ and $\downarrow$ electrons are shifted by $\pm\mu_{\rm B}H$, the change in their number is

$$\Delta N_\downarrow = -\Delta N_\uparrow = \mu_{\rm B}H\cdot\tfrac{1}{2}g(\varepsilon_{\rm F})\,V\ , \tag{4.70}$$

where $g(\varepsilon_{\rm F})$ is the density of states per unit volume (including both spin species), at the Fermi energy. Putting this all together, we find

$$M = \mu_{\rm B}^2\,g(\varepsilon_{\rm F})\,H \equiv \chi_{\rm P}\,H\ , \tag{4.71}$$

where $\chi_{\rm P} = \mu_{\rm B}^2\, g(\varepsilon_{\rm F})$ is the *Pauli susceptibility*. The Pauli susceptibility is positive, and hence is *paramagnetic*.

Using the formula for the density of states,

$$g(\varepsilon_{\rm F}) = \frac{m^* k_{\rm F}}{\pi^2 \hbar^2} \ , \tag{4.72}$$

we find

$$\chi_{\rm P} = \frac{1}{4\pi^2}\,\frac{m^*}{m}\left(\frac{e^2}{\hbar c}\right)^2 (k_{\rm F} a_{\rm B}) \ . \tag{4.73}$$

Using $e^2/\hbar c \simeq 1/137.036$ and assuming $k_{\rm F} a_{\rm B} \approx 1$, we find $\chi_{\rm P} \approx 10^{-6}$, which is comparable in magnitude (though opposite in sign) from the Larmor susceptibility of closed shells.

4.4.2 Landau Diamagnetism

Next, we investigate the orbital contribution. We assume a parabolic band, in which case

$$\mathcal{H} = \frac{1}{2m^*}\big(\boldsymbol{p} + \tfrac{e}{c}\boldsymbol{A}\big)^2 + \mu_{\rm B}\boldsymbol{\sigma}\cdot\boldsymbol{H} \ . \tag{4.74}$$

Appealing to the familiar results of a quantized charged particle in a uniform magnetic field, the energy levels are given by

$$\varepsilon(n, k_z, \sigma) = (n + \tfrac{1}{2})\,\hbar\omega_{\rm c} + \sigma\mu_{\rm B} H + \frac{\hbar^2 k_z^2}{2m^*} \ , \tag{4.75}$$

where $\omega_{\rm c} = eH/m^*c$ is the *cyclotron frequency*. Note that $\mu_{\rm B} H = \big(\frac{m^*}{m}\big)\cdot\frac{1}{2}\hbar\omega_{\rm c}$. The three-dimensional density of states is a convolution of the two-dimensional density of states,

$$g_{\rm 2d}(\varepsilon) = \frac{1}{2\pi\ell^2}\sum_{n=0}^{\infty}\delta\big(\varepsilon - (n+\tfrac{1}{2})\,\hbar\omega_{\rm c}\big) \ , \tag{4.76}$$

where $\ell = \sqrt{\hbar c/eH}$ is the magnetic length, and the one-dimensional density of states,

$$g_{\rm 1d}(\varepsilon) = \frac{1}{\pi}\frac{dk}{d\varepsilon} = \frac{\sqrt{m^*}}{\sqrt{2}\,\pi\hbar}\,\frac{1}{\sqrt{\varepsilon}} \ . \tag{4.77}$$

Thus,

$$g(\varepsilon) = \frac{\sqrt{m^*}}{\sqrt{2}\,\pi\hbar}\,\frac{1}{2\pi\ell^2}\sum_{n=0}^{\infty}\sum_{\sigma=\pm1}\frac{\Theta\big(\varepsilon-\varepsilon_{n\sigma}\big)}{\sqrt{\varepsilon-\varepsilon_{n\sigma}}} \ . \tag{4.78}$$

Thus, the grand potential,

$$\Omega(T,V,\mu,H) = -Vk_{\rm B}T\int\limits_{-\infty}^{\infty}\!\!d\varepsilon\, g(\varepsilon)\,\ln\Big\{1+e^{(\mu-\varepsilon)/k_{\rm B}T}\Big\} \tag{4.79}$$

may be written as the sum,

$$\Omega(T,V,\mu,H) = -Vk_{\rm B}T\frac{\sqrt{m^*}eH}{\sqrt{8}\,\pi^2\hbar^2c}\sum_{n=0}^{\infty}\sum_{\sigma=\pm1}F(\mu_\sigma - n\hbar\omega_{\rm c})\ , \tag{4.80}$$

with $\lambda = m^*/m$,

$$\mu_\sigma \equiv \mu - \tfrac{1}{2}(1+\sigma\lambda)\hbar\omega_{\rm c}\ , \tag{4.81}$$

and

$$F(\nu) = \int_0^\infty \frac{d\omega}{\sqrt{\omega}}\,\ln\left\{1+e^{(\nu-\omega)/k_{\rm B}T}\right\}\ . \tag{4.82}$$

We now invoke the Euler-MacLaurin formula,

$$\sum_{n=0}^{\infty} f(n) = \int_0^\infty dx\,f(x) + \tfrac{1}{2}f(0) - \tfrac{1}{12}\,f'(0) + \tfrac{1}{720}\,f'''(0) + \ldots\ , \tag{4.83}$$

which gives

$$\Omega = -\frac{Vk_{\rm B}T\,m^{*3/2}}{2\sqrt{2}\,\pi^2\hbar^3}\sum_{\sigma=\pm}\left\{\int_{-\infty}^{\mu_\sigma}\!\!d\varepsilon\,F(\varepsilon) + \tfrac{1}{2}\hbar\omega_{\rm c}\,F(\mu_\sigma) + \tfrac{1}{12}(\hbar\omega_{\rm c})^2F'(\mu_\sigma) + \ldots\right\}. \tag{4.84}$$

We now sum over σ and perform a Taylor expansion in $\hbar\omega_{\rm c} \propto H$, yielding

$$\begin{aligned}\Omega(T,V,\mu,H) &= -\frac{Vk_{\rm B}T\,m^{*3/2}}{2\sqrt{2}\,\pi^2\hbar^3}\sum_{\sigma=\pm}\left\{\int_{-\infty}^{\mu}\!\!d\varepsilon\,F(\varepsilon) + \tfrac{1}{8}\big(\lambda^2-\tfrac{1}{3}\big)(\hbar\omega_{\rm c})^2\,F'(\mu) + \mathcal{O}(H^4)\right\}\\ &= \left\{1+\tfrac{1}{2}\big(1-\tfrac{1}{3\lambda^2}\big)(\mu_{\rm B}H)^2\,\frac{\partial^2}{\partial\mu^2} + \mathcal{O}(H^4)\right\}\Omega(T,V,\mu,0)\ .\end{aligned} \tag{4.85}$$

Thus,

$$M = -\frac{1}{V}\big(1-\tfrac{1}{3\lambda^2}\big)\,\mu_{\rm B}^2\,H\,\frac{\partial^2\Omega}{\partial\mu^2}\bigg|_{H=0}\ , \tag{4.86}$$

and the zero field magnetic susceptibility is

$$\chi = \left(1-\frac{1}{3\lambda^2}\right)\mu_{\rm B}^2\,\frac{\partial n}{\partial\mu}\ . \tag{4.87}$$

The quantity $\chi_{\rm P} = \mu_{\rm B}^2\,(\partial n/\partial\mu)$ is simply the finite temperature Pauli susceptibility. The orbital contribution is negative, *i.e.* diamagnetic. Thus, $\chi = \chi_{\rm P} + \chi_{\rm L}$, where

$$\chi_{\rm L} = -\tfrac{1}{3}\,(m/m^*)^2\,\chi_{\rm P} \tag{4.88}$$

is the *Landau diamagnetic susceptibility*. For free electrons, $\lambda = m^*/m = 1$ and $\chi_{\rm L} = -\frac{1}{3}\chi_{\rm P}$ resulting in a reduced – but still paramagnetic – total susceptibility. However, in

semiconductors it is common to find $m^* \approx 0.1\,m$, in which case the Landau diamagnetic term overwhelms the Pauli paramagnetic term, and the overall susceptibility is negative.

In order to probe $\chi_{\rm P}$ without the diamagnetic $\chi_{\rm L}$ contribution, nuclear magnetic resonance (NMR) is used. NMR measures the effect of electron spins on the nuclear spins. The two are coupled via the hyperfine interaction,

$$\mathcal{H}_{\rm hf} = \frac{2g_{\rm N}\mu_{\rm B}\mu_{\rm N}}{\hbar^3\,r^3}\left\{\boldsymbol{L}\cdot\boldsymbol{I} - \boldsymbol{S}\cdot\boldsymbol{I} + (\hat{\boldsymbol{r}}\cdot\boldsymbol{S})(\hat{\boldsymbol{r}}\cdot\boldsymbol{I})\right\} + \frac{16\pi g_{\rm N}\mu_{\rm B}\mu_{\rm N}}{3\hbar}\,\boldsymbol{S}\cdot\boldsymbol{I}\,\delta(\boldsymbol{r})\ , \tag{4.89}$$

where $g_{\rm N}$ is the nuclear g-value, $\boldsymbol{I}$ is the nuclear total angular momentum, and $\mu_{\rm N}$ is the nuclear magneton.

4.5 Moment Formation in Interacting Itinerant Systems

4.5.1 The Hubbard Model

A noninteracting electron gas exhibits paramagnetism or diamagnetism, depending on the sign of χ, but never develops a spontaneous magnetic moment: $\boldsymbol{M}(\boldsymbol{H}=0)=0$. What gives rise to magnetism in solids? Overwhelmingly, the answer is that Coulomb repulsion between electrons is responsible for magnetism, in those instances in which magnetism arises. At first thought this might seem odd, since the Coulomb interaction is spin-independent. How then can it lead to a spontaneous magnetic moment?

To understand how Coulomb repulsion leads to magnetism, it is useful to consider a model interacting system, described by the Hamiltonian

$$\mathcal{H} = -t\sum_{\langle ij\rangle,\sigma}\left(c^\dagger_{i\sigma}c_{j\sigma} + c^\dagger_{j\sigma}c_{i\sigma}\right) + U\sum_i n_{i\uparrow}\,n_{i\downarrow} + \mu_{\rm B}\boldsymbol{H}\cdot\sum_{i,\alpha,\beta}c^\dagger_{i\alpha}\,\boldsymbol{\sigma}_{\alpha\beta}\,c_{i\beta}\ . \tag{4.90}$$

This is none other than the famous *Hubbard model*, which has served as a kind of Rosetta stone for interacting electron systems. The first term describes hopping of electrons along the links of some regular lattice (the symbol $\langle ij\rangle$ denotes a link between sites i and j). The second term describes the local (on-site) repulsion of electrons. This is a single orbital model, so the repulsion exists when one tries to put two electrons in the orbital, with opposite spin polarization. Typically the Hubbard U parameter is on the order of electron volts. The last term is the Zeeman interaction of the electron spins with an external magnetic field. Orbital effects can be modeled by associating a phase $\exp(iA_{ij})$ to the hopping matrix element t between sites i and j, where the directed sum of A_{ij} around a plaquette yields the total magnetic flux through the plaquette in units of $\phi_0 = hc/e$. We will ignore orbital effects here. Note that the interaction term is short-ranged, whereas the Coulomb interaction falls off as $1/|\boldsymbol{R}_i - \boldsymbol{R}_j|$. The Hubbard model is thus unrealistic, although screening effects in metals do effectively render the interaction to be short-ranged.

Within the Hubbard model, the interaction term is local and written as $Un_\uparrow n_\downarrow$ on any given site. This term favors a local moment. This is because the chemical potential will fix the

mean value of the total occupancy $n_\uparrow + n_\downarrow$, in which case it always pays to maximize the difference $|n_\uparrow - n_\downarrow|$.

4.5.2 Stoner Mean Field Theory

There are no general methods available to solve for even the ground state of an interacting many-body Hamiltonian. We'll solve this problem using a *mean field* theory due to Stoner. The idea is to write the occupancy $n_{i\sigma}$ as a sum of average and fluctuating terms:

$$n_{i\sigma} = \langle n_{i\sigma} \rangle + \delta n_{i\sigma} \ . \tag{4.91}$$

Here, $\langle n_{i\sigma} \rangle$ is the thermodynamic average; the above equation may then be taken as a definition of the fluctuating piece, $\delta n_{i\sigma}$. We assume that the average is site-independent. This is a significant assumption, for while we understand why each site should favor developing a moment, it is not clear that all these local moments should want to line up parallel to each other. Indeed, on a bipartite lattice, it is possible that the individual local moments on neighboring sites will be antiparallel, corresponding to an *antiferromagnetic* order of the pins. Our mean field theory will be one for *ferromagnetic* states.

We now write the interaction term as

$$\begin{aligned} n_{i\uparrow} n_{i\downarrow} &= \langle n_\uparrow \rangle \langle n_\downarrow \rangle + \langle n_\uparrow \rangle \, \delta n_{i\downarrow} + \langle n_\downarrow \rangle \, \delta n_{i\uparrow} + \overbrace{\delta n_{i\uparrow} \, \delta n_{i\downarrow}}^{(\text{fluct}^{\text{s}})^2} \\ &= -\langle n_\uparrow \rangle \langle n_\downarrow \rangle + \langle n_\uparrow \rangle \, n_{i\downarrow} + \langle n_\downarrow \rangle \, n_{i\uparrow} + \mathcal{O}\big((\delta n)^2\big) \\ &= \tfrac{1}{4}(m^2 - n^2) + \tfrac{1}{2} n \, (n_{i\uparrow} + n_{i\downarrow}) + \tfrac{1}{2} m \, (n_{i\uparrow} - n_{i\downarrow}) + \mathcal{O}\big((\delta n)^2\big) \ , \end{aligned} \tag{4.92}$$

where n and m are the average occupancy per spin and average spin polarization, each per unit cell:

$$n = \langle n_\downarrow \rangle + \langle n_\uparrow \rangle \tag{4.93}$$

$$m = \langle n_\downarrow \rangle - \langle n_\uparrow \rangle \ , \tag{4.94}$$

i.e. $\langle n_\sigma \rangle = \frac{1}{2}(n - \sigma m)$. The mean field grand canonical Hamiltonian $\mathcal{K} = \mathcal{H} - \mu \mathcal{N}$, may then be written as

$$\begin{aligned} \mathcal{K}^{\text{MF}} = &-\tfrac{1}{2} \sum_{i,j,\sigma} t_{ij} \Big(c^\dagger_{i\sigma} c_{j\sigma} + c^\dagger_{j\sigma} c_{i\sigma} \Big) - \big(\mu - \tfrac{1}{2} U n \big) \sum_{i\sigma} c^\dagger_{i\sigma} c_{i\sigma} \\ &+ \big(\mu_{\text{B}} H + \tfrac{1}{2} U m \big) \sum_{i\sigma} \sigma \, c^\dagger_{i\sigma} c_{i\sigma} + \tfrac{1}{4} N_{\text{sites}} \, U (m^2 - n^2) \ , \end{aligned} \tag{4.95}$$

where we've quantized spins along the direction of $\boldsymbol{H}$, defined as $\hat{\boldsymbol{z}}$. You should take note of two things here. First, the chemical potential is shifted *downward* (or the electron energies shifted *upward*) by an amount $\frac{1}{2}Un$, corresponding to the average energy of repulsion

with the background. Second, the effective magnetic field has been shifted by an amount $\frac{1}{2}Um/\mu_{\rm B}$, so the effective field is

$$H_{\rm eff} = H + \frac{Um}{2\mu_{\rm B}} \ . \tag{4.96}$$

The *bare* single particle dispersions are given by $\varepsilon_\sigma(\boldsymbol{k}) = -\hat{t}(\boldsymbol{k}) + \sigma\mu_{\rm B}H$, where

$$\hat{t}(\boldsymbol{k}) = \sum_{\boldsymbol{R}} t(\boldsymbol{R})\, e^{-i\boldsymbol{k}\cdot\boldsymbol{R}} \ , \tag{4.97}$$

and $t_{ij} = t(\boldsymbol{R}_i - \boldsymbol{R}_j)$. For nearest neighbor hopping on a d-dimensional cubic lattice, $\hat{t}(\boldsymbol{k}) = -t\sum_{\mu=1}^{d}\cos(k_\mu a)$, where a is the lattice constant. Including the mean field effects, the *effective* single particle dispersions become

$$\widetilde{\varepsilon}_\sigma(\boldsymbol{k}) = -\hat{t}(\boldsymbol{k}) - \tfrac{1}{2}Un + \big(\mu_{\rm B}H + \tfrac{1}{2}Um\big)\,\sigma \ . \tag{4.98}$$

We now solve the mean field theory, by obtaining the free energy per site, $\varphi(n,T,H)$. First, note that $\varphi = \omega + \mu n$, where $\omega = \Omega/N_{\rm sites}$ is the Landau, or grand canonical, free energy per site. This follows from the general relation $\Omega = F - \mu N$; note that the total electron number is $N = nN_{\rm sites}$, since n is the electron number per unit cell (including both spin species). If $g(\varepsilon)$ is the density of states per unit cell (rather than per unit volume), then we have[4]

$$\varphi = \tfrac{1}{4}U(m^2+n^2) + \bar{\mu}n - \tfrac{1}{2}k_{\rm B}T\int\limits_{-\infty}^{\infty}\!\!d\varepsilon\, g(\varepsilon)\left\{\ln\left(1+e^{(\bar{\mu}-\varepsilon-\Delta)/k_{\rm B}T}\right) + \ln\left(1+e^{(\bar{\mu}-\varepsilon+\Delta)/k_{\rm B}T}\right)\right\} \tag{4.99}$$

where $\bar{\mu} \equiv \mu - \frac{1}{2}Un$ and $\Delta \equiv \mu_{\rm B}H + \frac{1}{2}Um$. From this free energy we derive two self-consistent equations for μ and m. The first comes from demanding that φ be a function of n and not of μ, *i.e.* $\partial\varphi/\partial\mu = 0$, which leads to

$$n = \tfrac{1}{2}\int\limits_{-\infty}^{\infty}\!\!d\varepsilon\, g(\varepsilon)\Big\{f(\varepsilon-\Delta-\bar{\mu}) + f(\varepsilon+\Delta-\bar{\mu})\Big\} \ , \tag{4.100}$$

where $f(y) = \big[\exp(y/k_{\rm B}T)+1\big]^{-1}$ is the Fermi function. The second equation comes from minimizing f with respect to average moment m:

$$m = \tfrac{1}{2}\int\limits_{-\infty}^{\infty}\!\!d\varepsilon\, g(\varepsilon)\Big\{f(\varepsilon-\Delta-\bar{\mu}) - f(\varepsilon+\Delta-\bar{\mu})\Big\} \ . \tag{4.101}$$

Here, we will solve the first equation, eq. 4.100, and use the results to generate a Landau expansion of the free energy φ in powers of m^2. We assume that Δ is small, in which case

[4]Note that we have written $\mu n = \bar{\mu}n + \frac{1}{2}Un^2$, which explains the sign of the coefficient of n^2.

we may write

$$n = \int\limits_{-\infty}^{\infty} d\varepsilon\, g(\varepsilon) \left\{ f(\varepsilon - \bar{\mu}) + \tfrac{1}{2}\Delta^2\, f''(\varepsilon - \bar{\mu}) + \tfrac{1}{24}\,\Delta^4\, f''''(\varepsilon - \bar{\mu}) + \ldots \right\} . \tag{4.102}$$

We write $\bar{\mu}(\Delta) = \bar{\mu}_0 + \delta\bar{\mu}$ and expand in $\delta\bar{\mu}$. Since n is fixed in our (canonical) ensemble, we have

$$n = \int\limits_{-\infty}^{\infty} d\varepsilon\, g(\varepsilon)\, f\big(\varepsilon - \bar{\mu}_0\big) \ , \tag{4.103}$$

which defines $\bar{\mu}_0(n, T)$.[5] The remaining terms in the $\delta\bar{\mu}$ expansion of eqn. 4.102 must sum to zero. This yields

$$D(\bar{\mu}_0)\,\delta\bar{\mu} + \tfrac{1}{2}\Delta^2\, D'(\bar{\mu}_0) + \tfrac{1}{2}(\delta\bar{\mu})^2\, D'(\bar{\mu}_0) + \tfrac{1}{2}D''(\bar{\mu}_0)\,\Delta^2\,\delta\bar{\mu} + \tfrac{1}{24}\,D'''(\bar{\mu}_0)\,\Delta^4 + \mathcal{O}(\Delta^6) = 0 \ , \tag{4.104}$$

where

$$D(\mu) = -\int\limits_{-\infty}^{\infty} d\varepsilon\, g(\varepsilon)\, f'(\varepsilon - \mu) \tag{4.105}$$

is the thermally averaged *bare* density of states at energy μ. Note that the $k^{\rm th}$ derivative is

$$D^{(k)}(\mu) = -\int\limits_{-\infty}^{\infty} d\varepsilon\, g^{(k)}(\varepsilon)\, f'(\varepsilon - \mu) \ . \tag{4.106}$$

Solving for $\delta\bar{\mu}$, we obtain

$$\delta\bar{\mu} = -\tfrac{1}{2}a_1\Delta^2 - \tfrac{1}{24}\big(3a_1^3 - 6a_1a_2 + a_3\big)\,\Delta^4 + \mathcal{O}(\Delta^6) \ , \tag{4.107}$$

where

$$a_k \equiv \frac{D^{(k)}(\bar{\mu}_0)}{D(\bar{\mu}_0)} \ . \tag{4.108}$$

After integrating by parts and inserting this result for $\delta\bar{\mu}$ into our expression for the free energy f, we obtain the expansion

$$\varphi(n, T, m) = \varphi_0(n, T) + \tfrac{1}{4}Um^2 - \tfrac{1}{2}D(\bar{\mu}_0)\,\Delta^2 + \tfrac{1}{8}\left(\frac{\big[D'(\bar{\mu}_0)\big]^2}{D(\bar{\mu}_0)} - \tfrac{1}{3}\,D''(\bar{\mu}_0) \right) \Delta^4 + \ldots \ , \tag{4.109}$$

where prime denotes differentiation with respect to argument, at $m = 0$, and

$$\varphi_0(n, T) = \tfrac{1}{4}Un^2 + n\bar{\mu}_0 - \int\limits_{-\infty}^{\infty} d\varepsilon\, \mathcal{N}(\varepsilon)\, f\big(\varepsilon - \bar{\mu}_0\big) \ , \tag{4.110}$$

[5] The Gibbs-Duhem relation guarantees that such an equation of state exists, relating any three intensive thermodynamic quantities.

where $g(\varepsilon) = \mathcal{N}'(\varepsilon)$, so $\mathcal{N}(\varepsilon)$ is the integrated bare density of states per unit cell in the absence of any magnetic field (including both spin species).

We assume that H and m are small, in which case

$$\varphi = \varphi_0 + \tfrac{1}{2}am^2 + \tfrac{1}{4}bm^4 - \tfrac{1}{2}\chi_0\,H^2 - \frac{U\chi_0}{2\mu_{\rm B}}\,Hm + \dots \ , \tag{4.111}$$

where $\chi_0 = \mu_{\rm B}^2\,D(\bar\mu_0)$ is the Pauli susceptibility, and

$$a = \tfrac{1}{2}U\big(1 - \tfrac{1}{2}UD\big) \quad , \quad b = \tfrac{1}{32}\left(\frac{(D')^2}{D} - \tfrac{1}{3}\,D''\right)U^4 \quad , \tag{4.112}$$

where the argument of each $D^{(k)}$ above is $\bar\mu_0(n,T)$. The magnetization density (per unit cell, rather than per unit volume) is given by

$$M = -\frac{\partial\varphi}{\partial H} = \chi_0 H + \frac{U\chi_0}{2\mu_{\rm B}}\,m \ . \tag{4.113}$$

Minimizing with respect to m yields

$$am + bm^3 - \frac{U\chi_0}{2\mu_{\rm B}}\,H = 0 \ , \tag{4.114}$$

which gives, for small m,

$$m = \frac{\chi_0}{\mu_{\rm B}}\,\frac{H}{1 - \frac{1}{2}UD} \ . \tag{4.115}$$

We therefore obtain $M = \chi\,H$ with

$$\chi = \frac{\chi_0}{1 - \frac{U}{U_{\rm c}}} \ , \tag{4.116}$$

where

$$U_{\rm c} = \frac{2}{D(\bar\mu_0)} \ . \tag{4.117}$$

The denominator of χ increases the susceptibility above the bare Pauli value χ_0, and is referred to as – I kid you not – the *Stoner enhancement* (see Fig. 4.5).

It is worth emphasizing that the magnetization per unit cell is given by

$$M = -\frac{1}{N_{\rm sites}}\,\frac{\delta\mathcal{H}}{\delta H} = \mu_{\rm B}m \ . \tag{4.118}$$

This is an operator identity and is valid for any value of m, and not only small m.

When $H = 0$ we can still get a magnetic moment, provided $U > U_{\rm c}$. This is a consequence of the simple Landau theory we have derived. Solving for m when $H = 0$ gives $m = 0$ when $U < U_{\rm c}$ and

$$m(U) = \pm\left(\frac{U}{2b\,U_{\rm c}}\right)^{1/2}\sqrt{U - U_{\rm c}} \ , \tag{4.119}$$

when $U > U_{\rm c}$, and assuming $b > 0$. Thus we have the usual mean field order parameter exponent of $\beta = \frac{1}{2}$.

Figure 4.5: A graduate student experiences the Stoner enhancement.

4.5.3 Antiferromagnetic Solution

In addition to ferromagnetism, there may be other ordered states which solve the mean field theory. One such example is antiferromagnetism. On a bipartite lattice, the antiferromagnetic mean field theory is obtained from

$$\langle n_{i\sigma}\rangle = \tfrac{1}{2}n + \tfrac{1}{2}\sigma\, e^{i\boldsymbol{Q}\cdot\boldsymbol{R}_i}\, m \ , \tag{4.120}$$

where $\boldsymbol{Q} = (\pi/a, \pi/a, \ldots, \pi/a)$ is the antiferromagnetic ordering wavevector. The grand canonical Hamiltonian is then

$$\begin{aligned}\mathcal{K}^{\rm MF} &= -\tfrac{1}{2}\sum_{i,j,\sigma} t_{ij}\Big(c^\dagger_{i\sigma}c_{j\sigma} + c^\dagger_{j\sigma}c_{i\sigma}\Big) - \big(\mu - \tfrac{1}{2}Un\big)\sum_{i\sigma} c^\dagger_{i\sigma}c_{i\sigma} \\ &\quad + \tfrac{1}{2}Um\sum_{i\sigma} e^{i\boldsymbol{Q}\cdot\boldsymbol{R}_i}\,\sigma\, c^\dagger_{i\sigma}c_{i\sigma} + \tfrac{1}{4}N_{\rm sites}\,U(m^2-n^2) \qquad (4.121)\\ &= \tfrac{1}{2}\sum_{\boldsymbol{k}\sigma}\begin{pmatrix} c^\dagger_{\boldsymbol{k},\sigma} & c^\dagger_{\boldsymbol{k}+\boldsymbol{Q},\sigma}\end{pmatrix}\begin{pmatrix}\varepsilon(\boldsymbol{k})-\mu+\frac{1}{2}Un & \frac{1}{2}\sigma\,Um \\ \frac{1}{2}\sigma\,Um & \varepsilon(\boldsymbol{k}+\boldsymbol{Q})-\mu+\frac{1}{2}Un\end{pmatrix}\begin{pmatrix} c_{\boldsymbol{k},\sigma} \\ c_{\boldsymbol{k}+\boldsymbol{Q},\sigma}\end{pmatrix} \\ &\quad + \tfrac{1}{4}N_{\rm sites}\,U(m^2-n^2)\ , \qquad (4.122)\end{aligned}$$

where $\varepsilon(\boldsymbol{k}) = -\hat{t}(\boldsymbol{k})$, as before. On a bipartite lattice, with nearest neighbor hopping only, we have $\varepsilon(\boldsymbol{k}+\boldsymbol{Q}) = -\varepsilon(\boldsymbol{k})$. The above matrix is diagonalized by a unitary transformation, yielding the eigenvalues

$$\lambda_\pm = \pm\sqrt{\varepsilon^2(\boldsymbol{k}) + \Delta^2} - \bar{\mu} \tag{4.123}$$

with $\Delta = \frac{1}{2}Um$ and $\bar{\mu} = \mu - \frac{1}{2}Un$ as before. The free energy per unit cell is then

$$\begin{aligned}\varphi &= \tfrac{1}{4}U(m^2+n^2) + \bar{\mu}n \qquad (4.124)\\ &\quad - \tfrac{1}{2}k_{\rm B}T\int_{-\infty}^{\infty}\!d\varepsilon\, g(\varepsilon)\left\{\ln\left(1+e^{\left(\bar{\mu}-\sqrt{\varepsilon^2+\Delta^2}\right)/k_{\rm B}T}\right) + \ln\left(1+e^{\left(\bar{\mu}+\sqrt{\varepsilon^2+\Delta^2}\right)/k_{\rm B}T}\right)\right\}.\end{aligned}$$

The mean field equations are then

$$n = \tfrac{1}{2}\int_{-\infty}^{\infty} d\varepsilon\, g(\varepsilon)\left\{ f\Big(-\sqrt{\varepsilon^2+\Delta^2}-\bar{\mu}\Big) + f\Big(\sqrt{\varepsilon^2+\Delta^2}-\bar{\mu}\Big)\right\} \tag{4.125}$$

$$\frac{1}{U} = \tfrac{1}{2}\int_{-\infty}^{\infty} d\varepsilon\, \frac{g(\varepsilon)}{\sqrt{\varepsilon^2+\Delta^2}}\left\{ f\Big(-\sqrt{\varepsilon^2+\Delta^2}-\bar{\mu}\Big) - f\Big(\sqrt{\varepsilon^2+\Delta^2}-\bar{\mu}\Big)\right\}. \tag{4.126}$$

As in the case of the ferromagnet, a paramagnetic solution with $m = 0$ always exists, in which case the second of the above equations is no longer valid.

4.5.4 Mean Field Phase Diagram of the Hubbard Model

Let us compare the mean field theories for the ferromagnetic and antiferromagnetic states at $T = 0$ and $H = 0$. Due to particle-hole symmetry, we may assume $0 \le n \le 1$ without loss of generality. (The solutions repeat themselves under $n \to 2 - n$.) For the paramagnet, we have

$$n = \int_{-\infty}^{\bar{\mu}} d\varepsilon\, g(\varepsilon) \tag{4.127}$$

$$\varphi = \tfrac{1}{4}Un^2 + \int_{-\infty}^{\bar{\mu}} d\varepsilon\, g(\varepsilon)\,\varepsilon \quad , \tag{4.128}$$

with $\bar{\mu} = \mu - \frac{1}{2}Un$ is the 'renormalized' Fermi energy and $g(\varepsilon)$ is the density of states per unit cell in the absence of any explicit (H) or implicit (m) symmetry breaking, including both spin polarizations.

For the ferromagnet,

$$n = \tfrac{1}{2}\int_{-\infty}^{\bar{\mu}-\Delta} d\varepsilon\, g(\varepsilon) + \tfrac{1}{2}\int_{-\infty}^{\bar{\mu}+\Delta} d\varepsilon\, g(\varepsilon) \tag{4.129}$$

$$\frac{4\Delta}{U} = \int_{\bar{\mu}-\Delta}^{\bar{\mu}+\Delta} d\varepsilon\, g(\varepsilon) \tag{4.130}$$

$$\varphi = \tfrac{1}{4}Un^2 \;-\; \frac{\Delta^2}{U} \;+ \int_{-\infty}^{\bar{\mu}-\Delta} d\varepsilon\, g(\varepsilon)\,\varepsilon \;+ \int_{-\infty}^{\bar{\mu}+\Delta} d\varepsilon\, g(\varepsilon)\,\varepsilon \quad . \tag{4.131}$$

Here, $\Delta = \frac{1}{2}Um$ is nonzero in the ordered phase.

Finally, the antiferromagnetic mean field equations are

$$n_{\bar{\mu}<0} = \int_{\varepsilon_0}^{\infty} d\varepsilon\, g(\varepsilon) \qquad ; \qquad n_{\bar{\mu}>0} = 2 - \int_{\varepsilon_0}^{\infty} d\varepsilon\, g(\varepsilon) \tag{4.132}$$

$$\frac{2}{U} = \int_{\varepsilon_0}^{\infty} d\varepsilon\, \frac{g(\varepsilon)}{\sqrt{\varepsilon^2 + \Delta^2}} \tag{4.133}$$

$$\varphi = \tfrac{1}{4} U n^2 + \frac{\Delta^2}{U} - \int_{\varepsilon_0}^{\infty} d\varepsilon\, g(\varepsilon)\, \sqrt{\varepsilon^2 + \Delta^2} \quad , \tag{4.134}$$

where $\varepsilon_0 = \sqrt{\bar{\mu}^2 - \Delta^2}$ and $\Delta = \frac{1}{2} U m$ as before. Note that $|\bar{\mu}| \geq \Delta$ for these solutions. Exactly at half-filling, we have $n = 1$ and $\bar{\mu} = 0$. We then set $\varepsilon_0 = 0$.

The paramagnet to ferromagnet transition may be first or second order, depending on the details of $g(\varepsilon)$. If second order, it occurs at $U_{\rm c}^{\rm F} = 1\big/g(\bar{\mu}_{\rm P})$, where $\bar{\mu}_{\rm P}(n)$ is the paramagnetic solution for $\bar{\mu}$. The paramagnet to antiferromagnet transition is always second order in this mean field theory, since the RHS of eqn. (4.133) is a monotonic function of Δ. This transition occurs at $U_{\rm c}^{\rm A} = 2 \Big/ \int_{\bar{\mu}_{\rm P}}^{\infty} d\varepsilon\, g(\varepsilon)\, \varepsilon^{-1}$. Note that $U_{\rm c}^{\rm A} \to 0$ logarithmically for $n \to 1$, since $\bar{\mu}_{\rm P} = 0$ at half-filling.

For large U, the ferromagnetic solution always has the lowest energy, and therefore if $U_{\rm c}^{\rm A} < U_{\rm c}^{\rm F}$, there will be a first-order antiferromagnet to ferromagnet transition at some value $U^* > U_{\rm c}^{\rm F}$. In fig. 4.6, I plot the phase diagram obtained by solving the mean field equations assuming a semicircular density of states $g(\varepsilon) = \frac{2}{\pi} W^{-2} \sqrt{W^2 - \varepsilon^2}$. Also shown is the phase diagram for the $d = 2$ square lattice Hubbard model obtained by J. Hirsch (1985).

How well does Stoner theory describe the physics of the Hubbard model? Quantum Monte Carlo calculations by J. Hirsch (1985) found that the actual phase diagram of the $d = 2$ square lattice Hubbard Model exhibits no ferromagnetism for any n up to $U = 10$. Furthermore, he found the antiferromagnetic phase to be entirely confined to the vertical line $n = 1$. For $n \neq 1$ and $0 \leq U \leq 10$, the system is a paramagnet[6]. These results were state-of-the art at the time, but both computing power as well as numerical algorithms for interacting quantum systems have advanced considerably since 1985. Yet as of 2018, we *still* don't have a clear understanding of the $d = 2$ Hubbard model's $T = 0$ phase diagram! There is an emerging body of numerical evidence[7] that in the underdoped ($n < 1$) regime, there are portions of the phase diagram which exhibit a *stripe* ordering, in which antiferromagnetic order is interrupted by a parallel array of line defects containing excess holes (*i.e.* the absence of an electron)[8]. This problem has turned out to be unexpectedly rich, complex,

[6] A theorem due to Nagaoka establishes that the ground state is ferromagnetic for the case of a single hole in the $U = \infty$ system on bipartite lattices.

[7] See J. P. F. LeBlanc *et al.*, *Phys. Rev. X* **5**, 041041 (2015) and B. Zheng *et al.*, *Science* **358**, 1155 (2017).

[8] The best case for stripe order has been made at $T = 0$, $U/t = 8$, and hold doping $x = \frac{1}{8}$ (*i.e.* $n = \frac{7}{8}$).

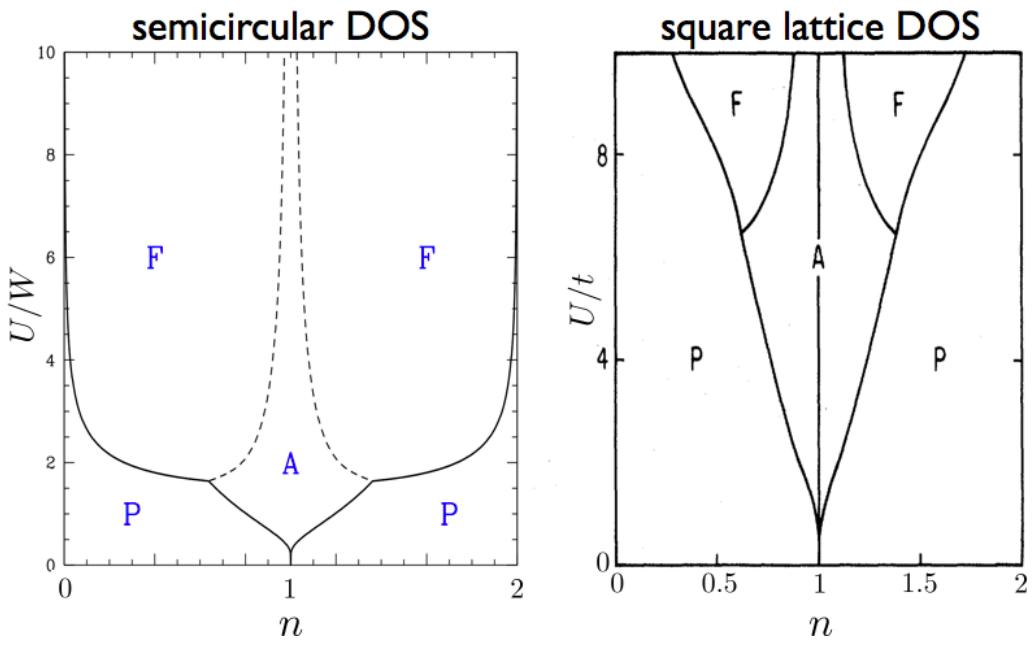

Figure 4.6: Mean field phase diagram of the Hubbard model, including paramagnetic (P), ferromagnetic (F), and antiferromagnetic (A) phases. Left panel: results using a semicircular density of states function of half-bandwidth W. Right panel: results using a two-dimensional square lattice density of states with nearest neighbor hopping t, from J. E. Hirsch, *Phys. Rev. B* **31**, 4403 (1985). The phase boundary between F and A phases is first order.

and numerically difficult to resolve due to the presence of *competing ordered states*, such as d-wave superconductivity and spiral magnetic phases, which lie nearby in energy with respect to the putative stripe ground state.

In order to achieve a ferromagnetic solution, it appears necessary to introduce geometric frustration, either by including a next-nearest-neighbor hopping amplitude t' or by defining the model on non-bipartite lattices. Numerical work by M. Ulmke (1997) showed the existence of a ferromagnetic phase at $T = 0$ on the FCC lattice Hubbard model for $U = 6$ and $n \in [0.15, 0.87]$ (approximately).

4.6 Interaction of Local Moments: the Heisenberg Model

While it is true that electrons have magnetic dipole moments, the corresponding dipole-dipole interactions in solids are usually negligible. This is easily seen by estimating the energy scale of the dipole-dipole interaction:

$$E_{\rm d-d} = \frac{\boldsymbol{m}_1 \cdot \boldsymbol{m}_2 - 3(\boldsymbol{m}_1 \cdot \hat{\boldsymbol{n}})(\boldsymbol{m}_2 \cdot \hat{\boldsymbol{n}})}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|^3} \ , \tag{4.135}$$

where $\hat{\boldsymbol{n}} = (\boldsymbol{r}_2 - \boldsymbol{r}_1)/|\boldsymbol{r}_2 - \boldsymbol{r}_1|$ is the direction vector pointing from $\boldsymbol{r}_1$ to $\boldsymbol{r}_2$. Substituting $\boldsymbol{m} = -\mu_{\rm B}\,\boldsymbol{\sigma}$, we estimate $E_{\rm d-d}$ as

$$|E_{\rm d-d}| \simeq \frac{\mu_{\rm B}^2}{R^3} = \frac{e^2}{2a_{\rm B}}\left(\frac{e^2}{\hbar c}\right)^2\left(\frac{a_{\rm B}}{R}\right)^3 , \tag{4.136}$$

and with $R \simeq 2.5$Å we obtain $E_{\rm d-d} \simeq 1\,\mu$eV, which is tiny on the scale of electronic energies in solids. The dominant magnetic coupling comes from the Coulomb interaction.

4.6.1 Ferromagnetic Exchange of Orthogonal Orbitals

In the Wannier basis, we may write the Coulomb interaction as

$$\hat{V} = \tfrac{1}{2}\sum_{\substack{\boldsymbol{R}_1,\boldsymbol{R}_2\\ \boldsymbol{R}_3,\boldsymbol{R}_4}}\sum_{\sigma,\sigma'}\big\langle\, \boldsymbol{R}_1\,\boldsymbol{R}_2\,\big|\,\frac{e^2}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\big|\,\boldsymbol{R}_4\,\boldsymbol{R}_3\,\big\rangle\, c^\dagger_{\boldsymbol{R}_1\sigma}\,c^\dagger_{\boldsymbol{R}_2\,\sigma'}\,c_{\boldsymbol{R}_3\,\sigma'}\,c_{\boldsymbol{R}_4\,\sigma}\ , \tag{4.137}$$

where we have assumed a single energy band. The Coulomb matrix element is

$$\big\langle\, \boldsymbol{R}_1\,\boldsymbol{R}_2\,\big|\,\frac{e^2}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\big|\,\boldsymbol{R}_4\,\boldsymbol{R}_3\,\big\rangle = \int\! d^3r\!\int\! d^3r'\,\varphi^*(\boldsymbol{r}-\boldsymbol{R}_1)\,\varphi^*(\boldsymbol{r}'-\boldsymbol{R}_2)\,\frac{e^2}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\varphi(\boldsymbol{r}'-\boldsymbol{R}_3)\,\varphi(\boldsymbol{r}-\boldsymbol{R}_4)\ . \tag{4.138}$$

Due to overlap factors, the matrix element will be small unless $\boldsymbol{R}_2 = \boldsymbol{R}_3$ and $\boldsymbol{R}_1 = \boldsymbol{R}_4$, in which case we obtain the *direct* Coulomb interaction,

$$\begin{aligned} V(\boldsymbol{R}-\boldsymbol{R}') &= \big\langle\, \boldsymbol{R}\,\boldsymbol{R}'\,\big|\,\frac{e^2}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\big|\,\boldsymbol{R}\,\boldsymbol{R}'\,\big\rangle \\ &= \int\! d^3r\!\int\! d^3r'\,\big|\varphi(\boldsymbol{r}-\boldsymbol{R})\big|^2\,\frac{e^2}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\big|\varphi(\boldsymbol{r}'-\boldsymbol{R}')\big|^2\ . \end{aligned} \tag{4.139}$$

The direct interaction decays as $|\boldsymbol{R}-\boldsymbol{R}'|^{-1}$ at large separations. Were we to include only these matrix elements, the second quantized form of the Coulomb interaction would be

$$\begin{aligned} \hat{V}_{\rm direct} &= \tfrac{1}{2}\sum_{\substack{\boldsymbol{R}\boldsymbol{R}'\\ \sigma\sigma'}} V(\boldsymbol{R}-\boldsymbol{R}')\Big(n_{\boldsymbol{R}\sigma}\,n_{\boldsymbol{R}'\sigma'} - \delta_{\boldsymbol{R}\boldsymbol{R}'}\,\delta_{\sigma\sigma'}\,n_{\boldsymbol{R}\sigma}\Big) \\ &= \sum_{\boldsymbol{R}} V(0)\,n_{\boldsymbol{R}\uparrow}\,n_{\boldsymbol{R}\downarrow} + \tfrac{1}{2}\sum_{\boldsymbol{R}\neq\boldsymbol{R}'} V(\boldsymbol{R}-\boldsymbol{R}')\,n_{\boldsymbol{R}}\,n_{\boldsymbol{R}'}\ , \end{aligned} \tag{4.140}$$

where $n_{\boldsymbol{R}} \equiv n_{\boldsymbol{R}\uparrow} + n_{\boldsymbol{R}\downarrow}$. The first term is the on-site Hubbard repulsion; one abbreviates $U \equiv V(0)$.

A second class of matrix elements can be identified: those with $\boldsymbol{R}_1 = \boldsymbol{R}_3 \equiv \boldsymbol{R}$ and $\boldsymbol{R}_2 = \boldsymbol{R}_4 \equiv \boldsymbol{R}'$, with $\boldsymbol{R} \neq \boldsymbol{R}'$. These are the so-called *exchange* integrals:

$$\begin{aligned} J(\boldsymbol{R}-\boldsymbol{R}') &= \big\langle\, \boldsymbol{R}\,\boldsymbol{R}'\,\big|\,\frac{e^2}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\big|\,\boldsymbol{R}'\,\boldsymbol{R}\,\big\rangle \\ &= \int\! d^3r\!\int\! d^3r'\,\varphi^*(\boldsymbol{r}-\boldsymbol{R})\,\varphi^*(\boldsymbol{r}'-\boldsymbol{R}')\,\frac{e^2}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\varphi(\boldsymbol{r}'-\boldsymbol{R})\,\varphi(\boldsymbol{r}-\boldsymbol{R}') \\ &= \int\! d^3r\!\int\! d^3r'\,\varphi^*(\boldsymbol{r})\,\varphi(\boldsymbol{r}+\boldsymbol{R}-\boldsymbol{R}')\,\frac{e^2}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\varphi^*(\boldsymbol{r}'+\boldsymbol{R}-\boldsymbol{R}')\,\varphi(\boldsymbol{r}')\ . \end{aligned} \tag{4.141}$$

Note that $J(\boldsymbol{R}-\boldsymbol{R}')$ is real. The exchange part of $\hat{V}$ is then

$$\begin{aligned}\hat{V}_{\text{exchange}} &= -\tfrac{1}{2}\sum_{\substack{\boldsymbol{R}\neq\boldsymbol{R}'\\ \sigma\sigma'}} J(\boldsymbol{R}-\boldsymbol{R}')\, c^\dagger_{\boldsymbol{R}\sigma}\, c_{\boldsymbol{R}\sigma'}\, c^\dagger_{\boldsymbol{R}'\sigma'}\, c_{\boldsymbol{R}'\sigma} \\ &= -\tfrac{1}{4}\sum_{\boldsymbol{R}\neq\boldsymbol{R}'} J(\boldsymbol{R}-\boldsymbol{R}')\left(n_{\boldsymbol{R}}\, n_{\boldsymbol{R}'} + \boldsymbol{\sigma}_{\boldsymbol{R}}\cdot\boldsymbol{\sigma}_{\boldsymbol{R}'}\right) .\end{aligned} \tag{4.142}$$

The $n_{\boldsymbol{R}}\, n_{\boldsymbol{R}'}$ piece can be lumped with the direct density-density interaction. What is new is the Heisenberg interaction,

$$\hat{V}_{\text{Heis}} = -\sum_{\boldsymbol{R}\neq\boldsymbol{R}'} J(\boldsymbol{R}-\boldsymbol{R}')\, \boldsymbol{S}_{\boldsymbol{R}}\cdot\boldsymbol{S}_{\boldsymbol{R}'} \ . \tag{4.143}$$

$J(\boldsymbol{R}-\boldsymbol{R}')$ is usually positive, and this gives us an explanation of Hund's first rule, which says to maximize S. This raises an interesting point, because we know that the ground state spatial wavefunction for the general two-body Hamiltonian

$$\mathcal{H} = -\frac{\hbar^2}{2m}\big(\nabla_1^2+\nabla_2^2\big) + V\big(|\boldsymbol{r}_1-\boldsymbol{r}_2|\big) \tag{4.144}$$

is nodeless. Thus, for fermions, the ground state spin wavefunction is an antisymmetric singlet state, corresponding to $S=0$. Yet the V^{3+} ion, with electronic configuration $[\text{Ar}]\,3d^2$, has a triplet $S=1$ ground state, according to Hund's first rule. Why don't the two 3d electrons have a singlet ground state, as the 'no nodes theorem' would seem to imply? The answer must have to do with the presence of the core electrons. Two electrons in the 1s shell do have a singlet ground state – indeed that is the only possibility. But the two 3d electrons in V^{3+} are not independent, but must be orthogonalized to the core states. This in effect projects out certain parts of the wavefunction, rendering the no nodes theorem inapplicable.

4.6.2 Heitler-London Theory of the H_2 Molecule

The Hamiltonian for the H_2 molecule is

$$\begin{aligned}\mathcal{H} = \frac{\boldsymbol{p}_1^2}{2m} - \frac{e^2}{|\boldsymbol{r}_1-\boldsymbol{R}_{\rm a}|} + \frac{\boldsymbol{p}_2^2}{2m} - \frac{e^2}{|\boldsymbol{r}_2-\boldsymbol{R}_{\rm b}|} \\ + \frac{e^2}{|\boldsymbol{R}_{\rm a}-\boldsymbol{R}_{\rm b}|} - \frac{e^2}{|\boldsymbol{r}_1-\boldsymbol{R}_{\rm b}|} - \frac{e^2}{|\boldsymbol{r}_2-\boldsymbol{R}_{\rm a}|} + \frac{e^2}{|\boldsymbol{r}_1-\boldsymbol{r}_2|} \ .\end{aligned} \tag{4.145}$$

The total wavefunction is antisymmetric: $\Psi(\boldsymbol{r}_1\,\sigma_1\,,\,\boldsymbol{r}_2\sigma_2\,) = -\Psi(\boldsymbol{r}_2\sigma_2\,,\,\boldsymbol{r}_1\sigma_1)$. The $N=2$ electron case is special because the wavefunction factorizes into a product:

$$\Psi(\boldsymbol{r}_1\,\sigma_1, \boldsymbol{r}_2\,\sigma_2) = \Phi(\boldsymbol{r}_1,\boldsymbol{r}_2)\,\chi(\sigma_1,\sigma_2) \ . \tag{4.146}$$

The spin wavefunction may either be symmetric (triplet, $S = 1$), or antisymmetric (singlet, $S = 0$):

$$\big|\,\chi\,\big\rangle = \begin{cases} \big|\uparrow\uparrow\big\rangle & S = 1 \\ \frac{1}{\sqrt{2}}\Big(\big|\uparrow\downarrow\big\rangle + \big|\downarrow\uparrow\big\rangle\Big) & S = 1 \\ \big|\downarrow\downarrow\big\rangle & S = 1 \\ \frac{1}{\sqrt{2}}\Big(\big|\uparrow\downarrow\big\rangle - \big|\downarrow\uparrow\big\rangle\Big) & S = 0\ . \end{cases} \tag{4.147}$$

A symmetric spin wavefunction requires an antisymmetric spatial one, and *vice versa*.

Despite the fact that $\mathcal{H}$ does not explicitly depend on spin, the effective low-energy Hamiltonian for this system is

$$\mathcal{H}_{\rm eff} = K + J\boldsymbol{S}_1 \cdot \boldsymbol{S}_2\ . \tag{4.148}$$

The singlet-triplet splitting is $\Delta E = E_{S=0} - E_{S=1} = -J$, so if $J > 0$ the ground state is the singlet, and if $J < 0$ the ground state is the three-fold degenerate triplet.

The one-electron 1s eigenfunction $\psi(\boldsymbol{r})$ satisfies the following eigenvalue equation:

$$\left\{-\frac{\hbar^2}{2m}\,\nabla^2 - \frac{e^2}{r}\right\}\psi(\boldsymbol{r}) = \varepsilon_0(\boldsymbol{r})\,\psi(\boldsymbol{r})\ . \tag{4.149}$$

In the Heitler-London approach, we write the two-electron wavefunction as a linear combination

$$\Phi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \alpha\,\Phi_{\rm I}(\boldsymbol{r}_1, \boldsymbol{r}_2) + \beta\,\Phi_{\rm II}(\boldsymbol{r}_1, \boldsymbol{r}_2)\ , \tag{4.150}$$

with

$$\begin{aligned} \Phi_{\rm I}(\boldsymbol{r}_1, \boldsymbol{r}_2) &= \psi(\boldsymbol{r}_1 - \boldsymbol{R}_{\rm a})\,\psi(\boldsymbol{r}_2 - \boldsymbol{R}_{\rm b}) \equiv \psi_{\rm a}(\boldsymbol{r}_1)\,\psi_{\rm b}(\boldsymbol{r}_2) \\ \Phi_{\rm II}(\boldsymbol{r}_1, \boldsymbol{r}_2) &= \psi(\boldsymbol{r}_1 - \boldsymbol{R}_{\rm b})\,\psi(\boldsymbol{r}_2 - \boldsymbol{R}_{\rm a}) \equiv \psi_{\rm b}(\boldsymbol{r}_1)\,\psi_{\rm a}(\boldsymbol{r}_2)\ . \end{aligned} \tag{4.151}$$

Assuming the atomic orbital $\psi(\boldsymbol{r})$ to be normalized, we define the following integrals:

$$\Delta = \int d^3r\,\psi_{\rm a}^*(\boldsymbol{r})\,\psi_{\rm b}(\boldsymbol{r}) \tag{4.152}$$

$$\begin{aligned} X &= \int d^3r_1 \int d^3r_2\,\big|\Phi_{\rm I}(\boldsymbol{r}_1, \boldsymbol{r}_2)\big|^2 \left(\frac{e^2}{R_{\rm ab}} + \frac{e^2}{r_{12}} - \frac{e^2}{r_{1\rm b}} - \frac{e^2}{r_{2\rm a}}\right) \\ &= \int d^3r_1 \int d^3r_2\,\big|\Phi_{\rm II}(\boldsymbol{r}_1, \boldsymbol{r}_2)\big|^2 \left(\frac{e^2}{R_{\rm ab}} + \frac{e^2}{r_{12}} - \frac{e^2}{r_{1\rm a}} - \frac{e^2}{r_{2\rm b}}\right) \end{aligned} \tag{4.153}$$

$$Y = \int d^3r_1 \int d^3r_2\,\Phi_{\rm I}^*(\boldsymbol{r}_1, \boldsymbol{r}_2)\,\Phi_{\rm II}(\boldsymbol{r}_1, \boldsymbol{r}_2)\left(\frac{e^2}{R_{\rm ab}} + \frac{e^2}{r_{12}} - \frac{e^2}{r_{1\rm b}} - \frac{e^2}{r_{2\rm a}}\right)\ , \tag{4.154}$$

with $\boldsymbol{r}_{1\rm a} = \boldsymbol{r}_1 - \boldsymbol{R}_{\rm a}$, *etc.* The expectation value of $\mathcal{H}$ in the state Φ is

$$\big\langle\,\Phi\,\big|\,\mathcal{H}\,\big|\,\Phi\,\big\rangle = (|\alpha|^2 + |\beta|^2)\,(2\varepsilon_0 + X) + (\alpha^*\beta + \beta^*\alpha)\,(2\varepsilon|\Delta|^2 + Y)\ , \tag{4.155}$$

and the self-overlap is

$$\left\langle \, \Phi \, \middle| \, \Phi \, \right\rangle = |\alpha|^2 + |\beta|^2 + |\Delta|^2 \, (\alpha^* \beta + \beta^* \alpha) \ . \tag{4.156}$$

We now minimize $\langle \mathcal{H} \rangle$ subject to the condition that Φ be normalized, using a Lagrange multiplier E to impose the normalization. Extremizing with respect to α^* and β^* yields

$$\begin{pmatrix} 2\varepsilon_0 + X & 2\varepsilon_0 |\Delta|^2 + Y \\ 2\varepsilon_0 |\Delta|^2 + Y & 2\varepsilon_0 + X \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E \begin{pmatrix} 1 & |\Delta|^2 \\ |\Delta|^2 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \ , \tag{4.157}$$

and extremizing with respect to E yields the normalization condition

$$|\alpha|^2 + |\beta|^2 + |\Delta|^2 \, (\alpha^* \beta + \beta^* \alpha) = 1 \ . \tag{4.158}$$

The solutions are symmetric and antisymmetric states, with $\beta/\alpha = \pm 1$, corresponding to the energies

$$E_\pm = 2\varepsilon_0 + \frac{X \pm Y}{1 \pm |\Delta|^2} \ . \tag{4.159}$$

Note that E_+ is the energy of the spatially symmetric state, which means a spin singlet while E_- corresponds to the spatially antisymmetric spin triplet.

The singlet-triplet splitting is

$$J = E_- - E_+ = 2 \, \frac{Y - X|\Delta|^2}{1 - |\Delta|^4} \ . \tag{4.160}$$

If $J > 0$, the triplet lies higher than the singlet, which says the ground state is antiferromagnetic. If $J < 0$, the triplet lies lower, and the ground state is ferromagnetic. The energy difference is largely determined by the Y integral:

$$Y = \int d^3 r_1 \int d^3 r_2 \, \Upsilon^*(\boldsymbol{r}_1) \, \Upsilon(\boldsymbol{r}_2) \left(\frac{e^2}{R_{\rm ab}} + \frac{e^2}{r_{12}} \right) - 2 \, \Delta^* \int d^3 r \, \psi_{\rm a}^*(\boldsymbol{r}) \frac{e^2}{|\boldsymbol{r} - \boldsymbol{R}_{\rm b}|} \, \psi_{\rm b}(\boldsymbol{r}) \ , \tag{4.161}$$

with $\Upsilon(\boldsymbol{r}) = \psi_{\rm a}^*(\boldsymbol{r}) \, \psi_{\rm b}(\boldsymbol{r})$. The first term is positive definite for the Coulomb interaction. The second term competes with the first if the overlap is considerable. The moral of the story now emerges:

$$\begin{aligned} \text{weak overlap} &\Longrightarrow \text{ferromagnetism } (J < 0) \\ \text{strong overlap} &\Longrightarrow \text{antiferromagnetism } (J > 0) \ . \end{aligned} \tag{4.162}$$

One finds that the H_2 molecule is indeed bound in the singlet state – the total energy has a minimum as a function of the separation $|\boldsymbol{R}_{\rm a} - \boldsymbol{R}_{\rm b}|$. In the triplet state, the molecule is unbound.

4.6.3 Failure of Heitler-London Theory

At large separations $R \equiv |\boldsymbol{R}_a - \boldsymbol{R}_b|$ the Heitler-London method describes two H atoms with tiny overlap of the electronic wavefunctions. But this tiny overlap is what determines

whether the ground state is a total spin singlet or triplet (we ignore coupling to the nuclear spin). Sugiura obtained the following expression for the singlet-triplet splitting in the $R \to \infty$ limit:

$$J(R) \simeq \left\{ \tfrac{56}{45} - \tfrac{4}{15}\gamma - \tfrac{4}{15} \ln\left(\frac{R}{a_{\rm H}}\right) \right\} \left(\frac{R}{a_{\rm H}}\right)^3 \left(\frac{e^2}{a_{\rm H}}\right) e^{-2R/a_{\rm H}} \ , \tag{4.163}$$

where $\gamma = 0.577\ldots$ is the Euler constant and where $\psi(r) = (\pi a_{\rm H}^3)^{-1/2} \exp(-r/a_{\rm H})$ is the hydrogenic wavefunction. This is negative for sufficiently large separations ($R > 50\, a_{\rm H}$). But this is a problem, since the eigenvalue problem is a Sturm-Liouville problem, hence the lowest energy eigenfunction must be spatially symmetric – *the singlet state must always lie at lower energy than the triplet*. The problem here is that Heitler-London theory does a good job on the wavefunction where it is large, *i.e.* in the vicinity of the protons, but a lousy job in the overlap region.

4.6.4 Herring's approach

Conyers Herring was the first to elucidate the failure of Heitler-London theory at large separations. He also showed how to properly derive a Heisenberg model for a lattice of hydrogenic orbitals. Herring started with the symmetric spatial wavefunction

$$\Psi(\boldsymbol{r}_1\ ,\ \ldots\ ,\ \boldsymbol{r}_N) = \prod_{i=1}^{N} \psi(\boldsymbol{r}_i - \boldsymbol{R}_i) \ . \tag{4.164}$$

This wavefunction would be appropriate were the electrons distinguishable. If we permute the electron coordinates using a spatial permutation $P_r \in \mathcal{S}_N$, we obtain another wavefunction of the same energy, E_0. However, there will be an overlap between these states:

$$J_P \equiv \big\langle\, \Psi \,\big|\, \mathcal{H} - E_0 \,\big|\, P_r\, \Psi \,\big\rangle \ . \tag{4.165}$$

The effective Hamiltonian is then

$$\mathcal{H}_{\rm eff} = E_0 + \sum_{P \in \mathcal{S}_N} J_P\, P_r \ . \tag{4.166}$$

A complete permutation P is a product of spatial and spin permutations: $P = P_r\, P_\sigma$, and the product when acting on an electronic wavefunction is $(-1)^P$, which is $+1$ for an even permutation and (-1) for an odd one[9]. Thus,

$$\mathcal{H}_{\rm eff} = E_0 + \sum_{P \in \mathcal{S}_N} (-1)^P\, J_P\, P_\sigma \ . \tag{4.167}$$

The spin permutation operators P_σ may be written in terms of the Pauli spin matrices, once we note that the two-cycle (ij) may be written

$$P_{(ij)} = \tfrac{1}{2} + \tfrac{1}{2}\, \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j \ . \tag{4.168}$$

[9]Here, 'even' and 'odd' refer to the number of 2-cycles into which a given permutation is decomposed.

Thus, accounting for only two-cycles, we have

$$\mathcal{H}_{\rm eff} = E_0 - \tfrac{1}{4}\sum_{i\neq j} J_{ij}\left(1+\boldsymbol{\sigma}_i\cdot\boldsymbol{\sigma}_j\right)\ . \tag{4.169}$$

For three-cycles, we have

$$\begin{aligned} P_{(ijk)} &= P_{(ik)}\,P_{(jk)} \\ &= \tfrac{1}{4}\big(1+\boldsymbol{\sigma}_i\cdot\boldsymbol{\sigma}_k\big)\big(1+\boldsymbol{\sigma}_j\cdot\boldsymbol{\sigma}_k\big) \\ &= \tfrac{1}{4}\Big[1+\boldsymbol{\sigma}_i\cdot\boldsymbol{\sigma}_j+\boldsymbol{\sigma}_j\cdot\boldsymbol{\sigma}_k+\boldsymbol{\sigma}_i\cdot\boldsymbol{\sigma}_k+i\boldsymbol{\sigma}_i\times\boldsymbol{\sigma}_j\cdot\boldsymbol{\sigma}_k\Big]\ . \end{aligned} \tag{4.170}$$

4.7 Mean Field Theory

We begin with the Heisenberg Hamiltonian

$$\mathcal{H} = -\tfrac{1}{2}\sum_{i,j} J_{ij}\,\boldsymbol{S}_i\cdot\boldsymbol{S}_j - \gamma\sum_i \boldsymbol{H}_i\cdot\boldsymbol{S}_i\ , \tag{4.171}$$

and write

$$\boldsymbol{S}_i = \boldsymbol{m}_i + \delta\boldsymbol{S}_i\ , \tag{4.172}$$

where $\boldsymbol{m}_i = \langle\boldsymbol{S}_i\rangle$ is the thermodynamic average of $\boldsymbol{S}_i$. We therefore have

$$\begin{aligned} \boldsymbol{S}_i\cdot\boldsymbol{S}_j &= \boldsymbol{m}_i\cdot\boldsymbol{m}_j + \boldsymbol{m}_i\cdot\delta\boldsymbol{S}_j + \boldsymbol{m}_j\cdot\delta\boldsymbol{S}_i + \delta\boldsymbol{S}_i\cdot\delta\boldsymbol{S}_j \\ &= -\boldsymbol{m}_i\cdot\boldsymbol{m}_j + \boldsymbol{m}_i\cdot\boldsymbol{S}_j + \boldsymbol{m}_j\cdot\boldsymbol{S}_i + \delta\boldsymbol{S}_i\cdot\delta\boldsymbol{S}_j\ . \end{aligned} \tag{4.173}$$

The last term is quadratic in the fluctuations, and as an approximation we ignore it. This results in the following *mean field Hamiltonian*,

$$\begin{aligned} \mathcal{H}^{\rm MF} &= +\tfrac{1}{2}\sum_{i,j} J_{ij}\,\boldsymbol{m}_i\cdot\boldsymbol{m}_j - \sum_i\Big(\gamma\boldsymbol{H}_i + \sum_j J_{ij}\,\boldsymbol{m}_j\Big)\cdot\boldsymbol{S}_i \\ &= E_0 - \gamma\sum_i \boldsymbol{H}_i^{\rm eff}\cdot\boldsymbol{S}_i\ , \end{aligned} \tag{4.174}$$

where

$$\begin{aligned} E_0 &= \tfrac{1}{2}\sum_{i,j} J_{ij}\,\boldsymbol{m}_i\cdot\boldsymbol{m}_j \\ \boldsymbol{H}_i^{\rm eff} &= \boldsymbol{H}_i + \gamma^{-1}\sum_j J_{ij}\,\boldsymbol{m}_j\ . \end{aligned} \tag{4.175}$$

Note how the *effective field* $\boldsymbol{H}_i^{\rm eff}$ is a sum of the external field $\boldsymbol{H}_i$ and the *internal field* $\boldsymbol{H}_i^{\rm int} = \gamma^{-1}\sum_j J_{ij}\,\boldsymbol{m}_j$. Self-consistency now requires that

$$\boldsymbol{m}_i = \frac{\mathsf{Tr}\ \boldsymbol{S}_i\exp\left(\gamma\boldsymbol{H}_i^{\rm eff}\cdot\boldsymbol{S}_i/k_{\rm B}T\right)}{\mathsf{Tr}\ \exp\left(\gamma\boldsymbol{H}_i^{\rm eff}\cdot\boldsymbol{S}_i/k_{\rm B}T\right)}\ , \tag{4.176}$$

where Tr means to sum or integrate over all local degrees of freedom (for site i). The free energy is then

$$F\big(\{\boldsymbol{m}_i\}\big) = \tfrac{1}{2}\sum_{i,j} J_{ij}\,\boldsymbol{m}_i\cdot\boldsymbol{m}_j - k_{\rm B}T\sum_i \ln \mathsf{Tr}\,\exp\left(\gamma \boldsymbol{H}_i^{\rm eff}\cdot\boldsymbol{S}_i/k_{\rm B}T\right)\,. \tag{4.177}$$

For classical systems, there are several common models:

- Ising Model with $S=\pm 1$:

$$\begin{aligned} m_i &= \tanh(\gamma H_i^{\rm eff}/k_{\rm B}T) \\ &= \tanh\Big(\beta\gamma H_i + \beta\sum_j J_{ij}\,m_j\Big)\,. \end{aligned} \tag{4.178}$$

  The free energy is

$$F = \tfrac{1}{2}\sum_{i,j} J_{ij}\,m_i\,m_j - k_{\rm B}T\sum_i \ln 2\cosh\Big(\beta\gamma H_i + \beta\sum_j J_{ij}\,m_j\Big)\,. \tag{4.179}$$

- Ising Model with $S=-1,0,+1$:

$$m_i = \frac{2\sinh\Big(\beta\gamma H_i + \beta\sum_j J_{ij}\,m_j\Big)}{1+2\cosh\Big(\beta\gamma H_i + \beta\sum_j J_{ij}\,m_j\Big)} \tag{4.180}$$

  and

$$F = \tfrac{1}{2}\sum_{i,j} J_{ij}\,m_i\,m_j - k_{\rm B}T\sum_i \ln\left\{1+2\cosh\Big(\beta\gamma H_i + \beta\sum_j J_{ij}\,m_j\Big)\right\}\,. \tag{4.181}$$

- XY model with $\boldsymbol{S}_i = (\cos\theta_i,\sin\theta_i)$, $\boldsymbol{H} = H\,\hat{\boldsymbol{x}}$

$$\begin{aligned} m_i = \big\langle\cos\theta_i\big\rangle &= \frac{\int_0^{2\pi} d\theta_i\,\cos\theta_i\,\exp\left(\gamma H_i^{\rm eff}\cos\theta_i/k_{\rm B}T\right)}{\int_0^{2\pi} d\theta_i\,\exp\left(\gamma H_i^{\rm eff}\cos\theta_i/k_{\rm B}T\right)} \\ &= \frac{I_1\Big(\beta\gamma H_i + \beta\sum_j J_{ij}\,m_j\Big)}{I_0\Big(\beta\gamma H_i + \beta\sum_j J_{ij}\,m_j\Big)}\,, \end{aligned} \tag{4.182}$$

  where $I_n(z)$ is a modified Bessel function. The free energy is

$$F = \tfrac{1}{2}\sum_{i,j} J_{ij}\,m_i\,m_j - k_{\rm B}T\sum_i \ln 2\pi I_0\Big(\beta\gamma H_i + \beta\sum_j J_{ij}\,m_j\Big)\,. \tag{4.183}$$

- O(3) model with $\boldsymbol{S}_i = (\sin\theta_i\cos\phi_i, \sin\theta_i\sin\phi_i, \cos\theta_i)$. Suppose that $\boldsymbol{m}_i$ points in the direction of $\boldsymbol{H}_i^{\rm eff}$. Then

$$\begin{aligned} m_i = \big\langle \cos\theta_i \big\rangle &= \frac{2\pi \int\limits_0^{2\pi} d\theta_i \, \sin\theta_i \, \cos\theta_i \, \exp\left(\gamma H_i^{\rm eff} \cos\theta_i / k_{\rm B}T\right)}{2\pi \int\limits_0^{2\pi} d\theta_i \, \sin\theta_i \, \exp\left(\gamma H_i^{\rm eff} \cos\theta_i / k_{\rm B}T\right)} \\ &= {\rm ctnh}\big(\gamma H_i^{\rm eff}/k_{\rm B}T\big) - \frac{k_{\rm B}T}{\gamma H_i^{\rm eff}} \\ &= {\rm ctnh}\Big(\beta\gamma H_i + \beta\sum_j J_{ij}\, m_j\Big) - \frac{k_{\rm B}T}{\gamma H_i + \sum_j J_{ij}\, m_j} \ . \end{aligned} \tag{4.184}$$

  The free energy is

$$F = \tfrac{1}{2}\sum_{i,j} J_{ij}\, m_i\, m_j - k_{\rm B}T \sum_i \ln\left\{ \frac{4\pi \sinh\Big(\beta\gamma H_i + \beta\sum_j J_{ij}\, m_j\Big)}{\beta\gamma H_i + \beta\sum_j J_{ij}\, m_j} \right\} . \tag{4.185}$$

<u>*EXERCISE:*</u> Show that the self-consistency is equivalent to $\partial F/\partial m_i = 0$.

4.7.1 Ferromagnets

Ising Model – Let us assume that the system orders ferromagnetically, with $m_i = m$ on all sites. Then, defining

$$\hat{J}(\boldsymbol{q}) = \sum_{\boldsymbol{R}} J(\boldsymbol{R})\, e^{-i\boldsymbol{q}\cdot\boldsymbol{R}} \ , \tag{4.186}$$

we have that the free energy per site, $f = F/N$, is

$$f(m) = \tfrac{1}{2}\hat{J}(0)\,\boldsymbol{m}^2 - k_{\rm B}T \ln {\rm Tr}\, \exp\Big\{ \big(\gamma\boldsymbol{H} + \hat{J}(0)\,\boldsymbol{m}\big)\cdot\boldsymbol{S}/k_{\rm B}T \Big\} \ . \tag{4.187}$$

For the $\mathbb{Z}_2$ (Ising) model, we have

$$m = \tanh\left(\beta\gamma H + \beta\hat{J}(0)m\right) \ , \tag{4.188}$$

a transcendental equation for m. For $H = 0$, we find $m = \tanh(\hat{J}(0)m/k_{\rm B}T)$, which yields the Curie temperature $T_{\rm C} = \hat{J}(0)/k_{\rm B}$.

O(3) Model – We have $\boldsymbol{m} = m\hat{\boldsymbol{H}}$ lies along $\boldsymbol{H}$. In the $\boldsymbol{H} \to 0$ limit, there is no preferred direction. The amplitude, however, satisfies

$$\frac{\partial f}{\partial m} = 0 \quad \Rightarrow \quad m = {\rm ctnh}\left(\hat{J}(0)\, m/k_{\rm B}T\right) - \frac{k_{\rm B}T}{\hat{J}(0)\, m} \ . \tag{4.189}$$

With $x \equiv \hat{J}(0)\, m/k_{\rm B}T$, then,

$$\frac{k_{\rm B}T}{\hat{J}(0)}\, x = {\rm ctnh}\, x - \frac{1}{x} = \frac{x}{3} - \frac{x^3}{45} + \ldots \ , \tag{4.190}$$

hence $T_{\rm c} = \hat{J}(0)/3k_{\rm B}$.

4.7.2 Antiferromagnets

If the lattice is bipartite, then we have *two* order parameters: $\boldsymbol{m}_{\rm A}$ and $\boldsymbol{m}_{\rm B}$. Suppose $J_{ij} = -J < 0$ if i and j are nearest neighbors, and zero otherwise. The effective fields on the A and B sublattices are given by

$$\boldsymbol{H}^{\rm eff}_{\rm A,B} \equiv \boldsymbol{H} - \gamma^{-1} z J \boldsymbol{m}_{\rm B,A} \ , \tag{4.191}$$

Note that the internal field on the A sublattice is $-\gamma^{-1} z J \boldsymbol{m}_{\rm B}$, while the internal field on the B sublattice is $-\gamma^{-1} z J \boldsymbol{m}_{\rm A}$. For the spin-$S$ quantum Heisenberg model, where $S^z \in \{-S, \ldots, +S\}$, we have

$$\mathsf{Tr}\, \exp(\boldsymbol{\xi} \cdot \boldsymbol{S}) = \frac{\sinh\left(S + \frac{1}{2}\right)\xi}{\sinh \frac{1}{2}\xi} \ , \tag{4.192}$$

hence, with $\boldsymbol{\xi} = \gamma \boldsymbol{H}^{\rm eff}_{\rm A,B}/k_{\rm B}T$, we have

$$\big\langle \boldsymbol{S} \big\rangle = \hat{\boldsymbol{\xi}}\, S\, B_S(S\xi) \tag{4.193}$$

where $B_S(x)$ is the Brillouin function,

$$B_S(x) = \Big(1 + \frac{1}{2S}\Big)\, \mathrm{ctnh}\, \Big(\big(1 + \frac{1}{2S}\big)\, x\Big) - \frac{1}{2S}\, \mathrm{ctnh}\, \Big(\frac{x}{2S}\Big) \ . \tag{4.194}$$

In order to best take advantage of the antiferromagnetic interaction and the external magnetic field, the ordered state is characterized by a *spin flop* in which $\boldsymbol{m}_{\rm A}$ and $\boldsymbol{m}_{\rm B}$ are, for weak fields, oriented in opposite directions in a plane perpendicular to $\boldsymbol{H}$, but each with a small component along $\boldsymbol{H}$.

When $H = 0$, the mean field equations take the form

$$\begin{aligned} m_{\rm A} &= S B_S\big(zJS m_{\rm B}/k_{\rm B}T\big) \\ m_{\rm B} &= S B_S\big(zJS m_{\rm A}/k_{\rm B}T\big) \ , \end{aligned} \tag{4.195}$$

where we have assumed $\boldsymbol{m}_{\rm A}$ and $\boldsymbol{m}_{\rm B}$ are antiparallel, with $\boldsymbol{m}_{\rm A} = m_{\rm A}\, \hat{\boldsymbol{n}}$ and $\boldsymbol{m}_{\rm B} = -m_{\rm B}\, \hat{\boldsymbol{n}}$, where $\hat{\boldsymbol{n}}$ is a unit vector. From the expansion of the Brillouin function, we obtain the Néel temperature $T_{\rm N} = zJ/k_{\rm B}$.

4.7.3 Susceptibility

For $T > T_{\rm c}$ the system is paramagnetic, and there is a linear response to an external field,

$$\begin{aligned} \chi^{\mu\nu}_{ij} &= \frac{\partial M^\mu_i}{\partial H^\nu_j} = \gamma\, \frac{\partial m^\mu_i}{\partial H^\nu_j} = -\frac{\partial^2 F}{\partial H^\mu_i\, \partial H^\nu_j} \\ &= \frac{\gamma^2}{k_{\rm B}T} \Big\{ \big\langle S^\mu_i\, S^\nu_j \big\rangle - \big\langle S^\mu_i \big\rangle \big\langle S^\nu_j \big\rangle \Big\} \end{aligned} \tag{4.196}$$

where $\{i, j\}$ are site indices and $\{\mu, \nu\}$ are internal spin indices. The mean field Hamiltonian is, up to a constant,

$$\mathcal{H}^{\rm MF} = -\gamma \sum_i \boldsymbol{H}_i^{\rm eff} \cdot \boldsymbol{S}_i \ , \tag{4.197}$$

which is a sum of single site terms. Hence, the response within $\mathcal{H}^{\rm MF}$ must be purely local as well as isotropic. That is, for weak effective fields, using $\boldsymbol{M}_i = \gamma \, \boldsymbol{m}_i$,

$$\boldsymbol{M}_i = \chi_0 \, \boldsymbol{H}_i^{\rm eff} = \chi_0 \, \boldsymbol{H}_i + \gamma^{-2} \, \chi_0 \, J_{ij} \, \boldsymbol{M}_j \ , \tag{4.198}$$

which is equivalent to

$$\big(\delta_{ij} - \gamma^{-2} \chi_0 \, J_{ij}\big) \boldsymbol{M}_j = \chi_0 \, \boldsymbol{H}_i \ , \tag{4.199}$$

and the mean field susceptibility is

$$\chi_{ij}^{\mu\nu} = \big[\chi_0^{-1} - \gamma^{-2} J\big]^{-1}_{ij} \, \delta^{\mu\nu} \ . \tag{4.200}$$

It is convenient to work in Fourier space, in which case the matrix structure is avoided and one has

$$\hat{\chi}(\boldsymbol{q}) = \frac{\chi_0}{1 - \gamma^{-2} \chi_0 \, \hat{J}(\boldsymbol{q})} \ . \tag{4.201}$$

The local susceptibility χ_0 is readily determined:

$$\begin{aligned} M^\mu = \gamma \left\langle S^\mu \right\rangle &= \gamma \, \frac{\mathsf{Tr} \ \boldsymbol{S} \, \exp(\gamma \boldsymbol{H} \cdot \boldsymbol{S} / k_{\rm B} T)}{\mathsf{Tr} \ \exp(\gamma \boldsymbol{H} \cdot \boldsymbol{S} / k_{\rm B} T)} \\ &= \gamma S B_S(S \gamma H / k_{\rm B} T) \, \hat{H}^\mu \ , \end{aligned} \tag{4.202}$$

where $B_S(x)$ is the Brillouin function from eqn. 4.194. As $\boldsymbol{H} \to 0$ we have $\boldsymbol{M} = \chi_0 \boldsymbol{H}$, with

$$\chi_0^{\mu\nu} = \frac{\gamma^2}{k_{\rm B} T} \cdot \frac{\mathsf{Tr} \, (S^\mu S^\nu)}{\mathsf{Tr} \ 1} \equiv \chi_0 \, \delta^{\mu\nu} \ , \tag{4.203}$$

where $\chi_0 = \frac{1}{N} \, \mathsf{Tr} \, (\boldsymbol{S}^2) / \, \mathsf{Tr} \ 1$, where N is the number of components of S^μ. Thus, for the Ising model ($N = 1$) we have $\chi_0^{\rm Ising} = \gamma^2 / k_{\rm B} T$, while for the spin-$S$ quantum Heisenberg model we have $\chi_0^{\rm Heis} = S(S+1)\gamma^2 / 3 k_{\rm B} T$. Note that $\chi_0 \propto T^{-1}$; the splitting of the degenerate energy levels by the magnetic field is of little consequence at high temperatures.

In many cases one deals with 'single ion anisotropy' terms. For example, one can add to the Heisenberg Hamiltonian a term such as

$$\mathcal{H}_{\rm a} = D \sum_i \big(S_i^z\big)^2 \ , \tag{4.204}$$

which for $D < 0$ results in an *easy axis* anisotropy (*i.e.* the spins prefer to align along the $\hat{\boldsymbol{z}}$-axis), and for $D > 0$ results in an *easy plane* anisotropy (*i.e.* the spins prefer to lie in the (x, y) plane). Since this term is already the sum of single site Hamiltonians, there is no need to subject it to a mean field treatment. One then obtains the mean field Hamiltonian

$$\mathcal{H}^{\rm MF} = D \sum_i \big(S_i^z\big)^2 - \gamma \sum_i \boldsymbol{H}_i^{\rm eff} \cdot \boldsymbol{S}_i \ . \tag{4.205}$$

In this case, χ_0 is now *anisotropic* in spin space. The general formula for χ_0 is

$$\chi_0^{\mu\nu} = \frac{\gamma^2}{k_{\rm B}T} \left\langle S^\mu S^\nu \right\rangle \tag{4.206}$$

where the thermodynamic average is taken with respect to the single site Hamiltonian.[10] One then has

$$\hat{\chi}_0^{\mu\nu}(\boldsymbol{q}) = \chi_0^{\mu\lambda} \left[\mathbb{I} - \gamma^{-2} \hat{J}(\boldsymbol{q}) \overset{\leftrightarrow}{\chi}_0 \right]^{-1}_{\lambda\nu} , \tag{4.207}$$

where the matrix inverse is now in internal spin space.

4.7.4 Variational Probability Distribution

Here's another way to derive mean field theory. Let Ω represent a configuration and let P_Ω be a probability distribution, normalized such that $\sum_\Omega P_\Omega = 1$. We define the entropy of the distribution as

$$S[P] = -k_{\rm B} \sum_\Omega P_\Omega \, \ln P_\Omega \ . \tag{4.208}$$

We now ask what distribution P_Ω minimizes the free energy $F = \langle \mathcal{H} \rangle - TS$. Working in an eigenbasis of $\mathcal{H}$, we have

$$F = \sum_\Omega P_\Omega \, E_\Omega + k_{\rm B}T \sum_\Omega P_\Omega \, \ln P_\Omega \ . \tag{4.209}$$

We extremize F subject to the normalization constraint, which is implemented with a Lagrange multiplier λ. This means we form the extended function

$$F^*\big(\{P_\Omega\}, \lambda\big) = \sum_\Omega P_\Omega \, E_\Omega + k_{\rm B}T \sum_\Omega P_\Omega \, \ln P_\Omega - \lambda\Big(\sum_\Omega P_\Omega - 1 \Big) \ , \tag{4.210}$$

and demand $dF^*/dP_\Omega = 0$ for all Ω as well as $dF^*/d\lambda = 0$. This results in the Boltzmann distribution,

$$P_\Omega^{\rm eq} = \frac{1}{Z} \, e^{-E_\Omega / k_{\rm B}T} \qquad , \qquad Z = \sum_l e^{-E_l / k_{\rm B}T} \ . \tag{4.211}$$

Thus, *any distribution other than* $P_\Omega^{\rm eq}$ *results in a larger free energy.*

Mean field theory may be formulated in terms of a variational probability distribution. Thus, rather than working with the Boltzmann distribution $P_\Omega^{\rm eq}$, which is usually intractable, we invoke a trial distribution $P_\Omega(x_1, x_2, \ldots)$, parameterized by $\{x_1, x_2, \ldots\}$, and minimize the resultant $F = \langle \mathcal{H} \rangle - TS$ with respect to those parameters.

As an example, consider the Ising model with spins $\sigma_i = \pm 1$. Each configuration is given by the set of spin polarizations: $\Omega = \{\sigma_1, \ldots, \sigma_N\}$. The full equilibrium probability distribution,

$$P_\Omega^{\rm eq} = Z^{-1} \, \exp\Big(\beta J \sum_{\langle ij \rangle} \sigma_i \sigma_j \Big) \ , \tag{4.212}$$

[10] Note that in (4.203) the single site Hamiltonian is simply $\mathcal{H}_0 = 0$.

with $\beta = 1/k_{\rm B}T$, is too cumbersome to work with. We replace this with a variational single-site distribution,

$$\begin{aligned} P_\Omega &= \prod_{j=1}^{N} P_i(\sigma_i) \\ P_i(\sigma_i) &= \tfrac{1}{2}(1+m_i)\,\delta_{\sigma_i,+1} + \tfrac{1}{2}(1-m_i)\,\delta_{\sigma_i,-1}\ . \end{aligned} \tag{4.213}$$

The variational parameters are $\{m_1, \ldots, m_N\}$. Note that P_Ω is properly normalized, by construction.

The entropy of our trial distribution is decomposed into a sum over single site terms:

$$\begin{aligned} S[P] &= \sum_i s(m_i) \\ s(m) &= -k_{\rm B}\left\{ \frac{1+m}{2}\,\ln\Big(\frac{1+m}{2}\Big) + \frac{1-m}{2}\,\ln\Big(\frac{1-m}{2}\Big)\right\} \end{aligned} \tag{4.214}$$

The thermodynamic average $\langle\sigma_i\rangle$ is simply

$$\langle\sigma_i\rangle = {\rm Tr}\, P_i(\sigma_i)\,\sigma_i = m_i\ , \tag{4.215}$$

hence from

$$\mathcal{H} = -\tfrac{1}{2}\sum_{i,j} J_{ij}\,\sigma_i\,\sigma_j - \gamma\sum_i H_i\,\sigma_i\ , \tag{4.216}$$

we derive the free energy

$$\begin{aligned} F\big(\{m_i\}\big) &= -\tfrac{1}{2}\sum_{i,j} J_{ij}\,m_i\,m_j - \gamma\sum_i H_i\,m_i \\ &\quad + k_{\rm B}T\sum_i\left\{ \frac{1+m_i}{2}\,\ln\Big(\frac{1+m_i}{2}\Big) + \frac{1-m_i}{2}\,\ln\Big(\frac{1-m_i}{2}\Big)\right\} \end{aligned} \tag{4.217}$$

Varying with respect to each m_i, we obtain the coupled nonlinear mean field equations,

$$m_i = \tanh\Big[\Big(\sum_j J_{ij}\,m_j + \gamma H_i\Big)/k_{\rm B}T\Big]\ . \tag{4.218}$$

For uniform magnetization ($m_i = m\ \forall\ i$), the free energy per site is

$$\begin{aligned} \frac{F}{N} &= -\tfrac{1}{2}\hat{J}(0)\,m^2 - \gamma H m + k_{\rm B}T\left\{ \frac{1+m}{2}\,\ln\Big(\frac{1+m}{2}\Big) + \frac{1-m}{2}\,\ln\Big(\frac{1-m}{2}\Big)\right\} \\ &= \tfrac{1}{2}\big(k_{\rm B}T - \hat{J}(0)\big)\,m^2 - \gamma H m + \tfrac{1}{12}\,k_{\rm B}T\,m^4 + \tfrac{1}{30}k_{\rm B}T m^6 + \ldots \end{aligned} \tag{4.219}$$

To compute the correlations, we may use the expression

$$\chi_{ij}(T) = \frac{\gamma^2}{k_{\rm B}T}\Big\{\langle\sigma_i\,\sigma_j\rangle - \langle\sigma_i\rangle\,\langle\sigma_j\rangle\Big\} \tag{4.220}$$

$$= \frac{\partial M_i}{\partial H_j} = \gamma\,\frac{\partial m_i}{\partial H_j} = -\frac{\partial^2 F}{\partial H_i\,\partial H_j}\ . \tag{4.221}$$

Thus, there are two ways to compute the susceptibility. One is to evaluate the spin-spin *correlation function*, as in (4.220). The other is to differentiate the magnetization to obtain the *response function*, as in (4.221). The equality between the two – called the "fluctuation-dissipation theorem" – is in fact only valid for the equilibrium Boltzmann distribution $P^{\rm eq}_\Omega$. Which side of the equation should we use in our variational mean field theory? It is more accurate to use the response function. To roughly see this, let us write $P = P^{\rm eq} + \delta P$, with δP small in some sense. The free energy is given by

$$F[P] = F[P^{\rm eq}] + \delta P \cdot \frac{\delta F}{\delta P}\bigg|_{P=P^{\rm eq}} + \mathcal{O}\big((\delta P)^2\big) \ . \tag{4.222}$$

Our variational treatment guarantees that the second term vanishes, since we extremize F with respect to P. Thus, in some sense, the error in F is only of order $(\delta P)^2$. If we compute the correlation function using $\langle A\rangle = {\rm Tr}\,(P\,A)$, where A is any operator, then the error will be linear in δP. So it is better to use the response function than the correlation function.

EXERCISE: Articulate the correspondence between this variational version of mean field theory and the 'neglect of fluctuations' approach we derived earlier.

4.8 Magnetic Ordering

The $\boldsymbol{q}$-dependent susceptibility in (4.201) diverges when $\gamma^{-2}\chi_0\hat{J}(\boldsymbol{q}) = 1$. As we know, such a divergence heralds the onset of a *phase transition* where there is a *spontaneous magnetization* in the ordered (*i.e.* low temperature) phase. Typically this happens at a particular wavevector $\boldsymbol{Q}$, or a set of symmetry related wavevectors $\{\boldsymbol{Q}_1, \boldsymbol{Q}_2, \ldots\}$. The ordering wavevector is that value of $\boldsymbol{q}$ which results in a *maximum* of $\hat{J}(\boldsymbol{q})$: $\max_{\boldsymbol{q}}\big\{\hat{J}(\boldsymbol{q})\big\} = \hat{J}(\boldsymbol{Q})$. The susceptibility, for isotropic systems, can be written

$$\hat{\chi}(\boldsymbol{q}) = \frac{\chi_0}{\big[1 - \gamma^{-2}\chi_0\,\hat{J}(\boldsymbol{Q})\big] + \gamma^{-2}\chi_0\big[\hat{J}(\boldsymbol{Q}) - \hat{J}(\boldsymbol{q})\big]} \ . \tag{4.223}$$

The critical temperature $T_{\rm c}$ is determined by the relation

$$\gamma^{-2}\chi_0(T_{\rm c})\,\hat{J}(\boldsymbol{Q}) = 1. \tag{4.224}$$

Expanding about $T = T_{\rm c}$, and about $\boldsymbol{q} = \boldsymbol{Q}$, where

$$\hat{J}(\boldsymbol{q}) = \hat{J}(\boldsymbol{Q})\Big\{1 - (\boldsymbol{q}-\boldsymbol{Q})^2 R_*^2 + \ldots\Big\} \ , \tag{4.225}$$

we have

$$\hat{\chi}(\boldsymbol{q}) \approx \frac{\chi_0/R_*^2}{\xi^{-2}(T) + (\boldsymbol{q}-\boldsymbol{Q})^2} \ , \tag{4.226}$$

where

$$\xi^{-2}(T) = -\frac{\chi_0'(T_{\rm c})}{\chi_0(T_{\rm c})} \cdot R_*^{-2} \cdot (T - T_{\rm c}) \ . \tag{4.227}$$

Thus, $\xi(T) \propto (T - T_{\rm c})^{-1/2}$. The real space susceptibility $\chi(\boldsymbol{R}_i - \boldsymbol{R}_j)$ oscillates with wavevector $\boldsymbol{Q}$ and decays on the scale of the correlation length $\xi(T)$.

- *Ferromagnet:* $J_{ij} = +J > 0$ if i and j are nearest neighbors; otherwise $J_{ij} = 0$. On a hypercubic lattice (d dimensions, $2d$ nearest neighbors), we then have

$$\hat{J}(\boldsymbol{q}) = J \sum_{\boldsymbol{\delta}} e^{-i\boldsymbol{q}\cdot\boldsymbol{\delta}} = 2J\big\{\cos(q_1 a) + \cos(q_2 a) + \ldots + \cos(q_d a)\big\} \ . \tag{4.228}$$

  The ordering wavevector is $\boldsymbol{Q} = 0$, and $\hat{J}(\boldsymbol{Q}) = 2dJ$. For the spin-$S$ Heisenberg model, then, $T_{\rm C} = \frac{2}{3} d\, S(S+1)\, J/k_{\rm B}$, and the susceptibility is

$$\hat{\chi}(\boldsymbol{q}) = \frac{\frac{1}{3}\gamma^2 S(S+1)/k_{\rm B}}{(T - T_{\rm C}) + T_{\rm C}\, d^{-1} \sum_{\nu=1}^{d} \big[1 - \cos(q_\nu a)\big]} \ . \tag{4.229}$$

  The uniform susceptibility $\chi = \hat{\chi}(\boldsymbol{q} = 0)$ is then

$$\chi(T) = \frac{\gamma^2 S(S+1)}{3k_{\rm B}(T - T_{\rm C})} \ . \tag{4.230}$$

  Ferromagnetic insulators: ferrites, EuO, TDAE-C_{60}.

- *Antiferromagnet:* $J_{ij} = -J < 0$ if i and j are nearest neighbors; otherwise $J_{ij} = 0$. On a hypercubic lattice (d dimensions, $2d$ nearest neighbors), we then have

$$\hat{J}(\boldsymbol{q}) = -J \sum_{\boldsymbol{\delta}} e^{-i\boldsymbol{q}\cdot\boldsymbol{\delta}} = -2J\big\{\cos(q_1 a) + \cos(q_2 a) + \ldots + \cos(q_d a)\big\} \ . \tag{4.231}$$

  The ordering wavevector is $\boldsymbol{Q} = (\pi/a, \ldots, \pi/a)$, at the zone corner, where $\hat{J}(\boldsymbol{Q}) = 2dJ$. For the spin-$S$ Heisenberg model, then, $T_{\rm N} = \frac{2}{3} d\, S(S+1)\, J/k_{\rm B}$, and the susceptibility is

$$\hat{\chi}(\boldsymbol{q}) = \frac{\gamma^2 S(S+1)/3k_{\rm B}}{(T - T_{\rm N}) + T_{\rm N}\, d^{-1} \sum_{\nu=1}^{d} \big[1 + \cos(q_\nu a)\big]} \ . \tag{4.232}$$

  The uniform susceptibility $\chi = \hat{\chi}(\boldsymbol{q} = 0)$ is then

$$\chi(T) = \frac{\gamma^2 S(S+1)}{3k_{\rm B}(T + T_{\rm N})} \ , \tag{4.233}$$

  which *does not diverge*. Indeed, plotting $\chi^{-1}(T)$ *versus* T, one obtains an intercept along the T-axis at $T = -T_{\rm N}$. This is one crude way of estimating the Néel temperature. What *does* diverge is the *staggered susceptibility* $\chi_{\rm stag} \equiv \hat{\chi}(\boldsymbol{Q}, T)$, *i.e.* the susceptibility at the ordering wavevector:

$$\chi_{\rm stag}(T) = \frac{\gamma^2 S(S+1)}{3k_{\rm B}(T - T_{\rm N})} \ . \tag{4.234}$$

- *Frustrated Antiferromagnet:* On the triangular lattice, the antiferromagnetic state is frustrated. What does mean field theory predict? We begin by writing primitive

direct lattice vectors $\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$ and primitive reciprocal lattice vectors $\{\boldsymbol{b}_1, \boldsymbol{b}_2\}$:

$$\boldsymbol{a}_1 = a\,(1, 0) \qquad\qquad \boldsymbol{b}_1 = \frac{4\pi}{a\sqrt{3}}\,\big(\tfrac{\sqrt{3}}{2}, -\tfrac{1}{2}\big) \tag{4.235}$$

$$\boldsymbol{a}_2 = a\,\big(\tfrac{1}{2}, \tfrac{\sqrt{3}}{2}\big) \qquad\qquad \boldsymbol{b}_2 = \frac{4\pi}{a\sqrt{3}}\,(0, 1)\ , \tag{4.236}$$

where a is the lattice constant. The six nearest neighbor vectors are then

$$\delta \in \Big\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_2 - \boldsymbol{a}_1, -\boldsymbol{a}_1, -\boldsymbol{a}_2, \boldsymbol{a}_1 - \boldsymbol{a}_2\Big\}\ , \tag{4.237}$$

and writing $\boldsymbol{q} \equiv x_1\boldsymbol{b}_1 + x_2\boldsymbol{b}_2$, we find

$$\hat{J}(\boldsymbol{q}) = -2J\big\{\cos(2\pi x_1) + \cos(2\pi x_2) + \cos(2\pi x_1 - 2\pi x_2)\big\}\ . \tag{4.238}$$

We suspect that this should be maximized somewhere along the perimeter of the Brillouin zone. The face center lies at $(x_1, x_2) = (\frac{1}{2}, \frac{1}{2})$, where $\hat{J}(\boldsymbol{q}) = +2J$. However, an even greater value is obtained either of the two inequivalent zone corners, $(x_1, x_2) = (\frac{2}{3}, \frac{1}{3})$ and $(x_1, x_2) = (\frac{1}{3}, \frac{2}{3})$, where $\hat{J}(\boldsymbol{q}) = +3J$. Each of these corresponds to a *tripartite* division of the triangular lattice in to three $\sqrt{3} \times \sqrt{3}$ triangular sublattices.

Antiferromagnetic insulators: MnO, CoO, FeO, NiO, La_2CuO_4.

- *Helimagnet:* Consider a cubic lattice system with mixed ferromagnetic and antiferromagnetic interactions:

$$J_{ij} = \begin{cases} +J_1 > 0 & \text{6 nearest neighbors} \\ -J_2 < 0 & \text{12 next-nearest neighbors} \\ 0 & \text{otherwise}\ . \end{cases} \tag{4.239}$$

Then

$$\begin{aligned}\hat{J}(\boldsymbol{q}) = {} & 2J_1\big[\cos(q_x a) + \cos(q_y a) + \cos(q_z a)\big] \\ & - 4J_2\big[\cos(q_x a)\cos(q_y a) + \cos(q_x a)\cos(q_z a) + \cos(q_y a)\cos(q_z a)\big]\ .\end{aligned} \tag{4.240}$$

The ordering wavevector is then

$$\boldsymbol{Q} = \begin{cases} a^{-1}\cos^{-1}\left(\frac{J_1}{4J_2}\right)(\hat{\boldsymbol{x}} + \hat{\boldsymbol{y}} + \hat{\boldsymbol{z}}) & \text{if } J_1 < 4J_2 \\ 0 & \text{if } J_1 \geq 4J_2\ . \end{cases} \tag{4.241}$$

Thus, for $J_1 < 4J_2$ the order is *incommensurate* with the lattice. The maximum value of $\hat{J}(\boldsymbol{q})$ is

$$\hat{J}(\boldsymbol{Q}) = \begin{cases} \frac{3J_1^2}{4J_2} & \text{if } J_1 < 4J_2 \\ 6(J_1 - 2J_2) & \text{if } J_1 \geq 4J_2\ , \end{cases} \tag{4.242}$$

hence incommensurate order sets in at $T_{\rm I} = S(S+1)J_1^2/4k_{\rm B}J_2$. The uniform susceptibility is

$$\hat{\chi}(0) = \frac{\gamma^2 S(S+1)/3k_{\rm B}}{T - 8\,T_{\rm I}\,\frac{J_2}{J_1}\Big(1 - \frac{2J_2}{J_1}\Big)}\ . \tag{4.243}$$

Thus,

$$\chi(T) \simeq \begin{cases} \frac{\mathcal{C}}{T+T^*} & 0 < J_1 < 2J_2 \text{ (like AFM)} \\ \\ \frac{\mathcal{C}}{T-T^*} & 2J_2 < J_1 < 4J_2 \text{ (like FM) .} \end{cases} \tag{4.244}$$

4.8.1 Mean Field Theory of Anisotropic Magnetic Systems

Consider the anisotropic Heisenberg model,

$$\mathcal{H} = -\overbrace{\sum_{i<j} J_{ij}^{\parallel} \boldsymbol{S}_i \cdot \boldsymbol{S}_j}^{intra} - \overbrace{\sum_{i<j} J_{ij}^{\perp} \boldsymbol{S}_i \cdot \boldsymbol{S}_j}^{inter} - \gamma \sum_i \boldsymbol{H}_i \cdot \boldsymbol{S}_i \ . \tag{4.245}$$

Here, $J_{ij}^{\parallel}$ only connects sites within the same plane (quasi-2d) or chain (quasi-1d), while $J_{ij}^{\perp}$ only connects sites in different planes/chains. We assume that we have an adequate theory for isolated plains/chains, and we effect a mean field decomposition on the inter-plane/interchain term:

$$\boldsymbol{S}_i \cdot \boldsymbol{S}_j = -\langle \boldsymbol{S}_i \rangle \cdot \langle \boldsymbol{S}_j \rangle + \langle \boldsymbol{S}_i \rangle \cdot \boldsymbol{S}_j + \langle \boldsymbol{S}_j \rangle \cdot \boldsymbol{S}_i + \overbrace{\delta \boldsymbol{S}_i \cdot \delta \boldsymbol{S}_j}^{(\text{fluct})^2} \ , \tag{4.246}$$

resulting in the effective field

$$\boldsymbol{H}^{\text{eff}}(\boldsymbol{q}, \omega) = \boldsymbol{H}(\boldsymbol{q}, \omega) + \gamma^{-2} \hat{J}^{\perp}(\boldsymbol{q}_{\perp}) \, \boldsymbol{M}(\boldsymbol{q}, \omega) \ , \tag{4.247}$$

where $\boldsymbol{M}(\boldsymbol{q}, \omega) = \gamma \langle \boldsymbol{S}(\boldsymbol{q}, \omega) \rangle$. Thus,

$$\hat{\chi}(\boldsymbol{q}, \omega) = \frac{\hat{\chi}^{\parallel}(\boldsymbol{q}_{\parallel}, \omega)}{1 - \gamma^{-2} \hat{J}^{\perp}(\boldsymbol{q}_{\perp}) \, \hat{\chi}^{\parallel}(\boldsymbol{q}_{\parallel}, \omega)} \ , \tag{4.248}$$

where $\hat{\chi}^{\parallel}(\boldsymbol{q}_{\parallel}, \omega)$ is assumed known.

4.8.2 Quasi-1D Chains

Consider a ferromagnet on a cubic lattice where the exchange interaction along the $\hat{\boldsymbol{z}}$-direction ($\parallel$) is much larger than that in the (x, y) plane ($\perp$). Treating the in-plane interactions via mean field theory, we have

$$\hat{\chi}(\boldsymbol{q}_{\perp}, q_z) = \frac{\hat{\chi}_{1\text{D}}(q_z)}{1 - \gamma^{-2} \hat{J}^{\perp}(\boldsymbol{q}_{\perp}) \, \hat{\chi}_{1\text{D}}(q_z)} \ , \tag{4.249}$$

with

$$\hat{J}^{\perp}(\boldsymbol{q}_{\perp}) = 2J_{\perp} \Big\{ \cos(q_x a) + \cos(q_y a) \Big\} \ . \tag{4.250}$$

For the Ising model we can compute $\hat\chi_{\rm 1D}(q_z)$ exactly using the high temperature expansion:

$$\begin{aligned}\langle\sigma_n\,\sigma_{n'}\rangle &= \frac{\mathsf{Tr}\,\Big\{\sigma_n\,\sigma_{n'}\,\prod_j\big(1+\tanh(J_\parallel/k_{\rm B}T)\,\sigma_j\,\sigma_{j+1}\big)\Big\}}{\mathsf{Tr}\,\prod_j\big(1+\tanh(J_\parallel/k_{\rm B}T)\,\sigma_j\,\sigma_{j+1}\big)} \\ &= \tanh^{|n-n'|}(J_\parallel/k_{\rm B}T)\ .\end{aligned} \tag{4.251}$$

Thus,

$$\begin{aligned}\hat\chi_{\rm 1D}(q_z) &= \frac{\gamma^2}{k_{\rm B}T}\sum_{n=-\infty}^{\infty}\tanh^{|n|}(J_\parallel/k_{\rm B}T)\,e^{inq_zc} \\ &= \frac{\gamma^2}{k_{\rm B}T}\,\frac{1}{\cosh(2J_\parallel/k_{\rm B}T)-\sinh(2J_\parallel/k_{\rm B}T)\,\cos(q_zc)} \\ &\approx \frac{2\pi\gamma^2}{ck_{\rm B}T}\cdot\frac{1}{\pi}\,\frac{\xi^{-1}}{\xi^{-2}+q_z^2}\ ,\end{aligned} \tag{4.252}$$

where c is the lattice spacing along the chains, and where the last approximation is valid for $q\to0$ and $\xi\to\infty$. The correlation length in this limit is given by

$$\xi(T)\simeq\frac{c}{2}\,\exp(2J_\parallel/k_{\rm B}T)\ . \tag{4.253}$$

Note that $\xi(T)$ diverges only at $T=0$. This is consistent with the well-known fact that the *lower critical dimension* for systems with *discrete global symmetries* and *short-ranged interactions* is $d=1$. That is to say that there is no spontaneous breaking of any discrete symmetry in one-dimension (with the proviso of sufficiently short-ranged interactions). For *continuous symmetries* the lower critical dimension is $d=2$, which is the content of the Hohenberg-Mermin-Wagner (HMW) theorem.

Accounting for the residual interchain interactions via mean field theory, we obtain the anisotropic (in space) susceptibility

$$\hat\chi(\boldsymbol{q}_\perp,q_z)=\frac{\hat\chi_{\rm 1D}(q_z)}{1-\gamma^{-2}\cdot2J_\perp\big\{\cos(q_xa)+\cos(q_ya)\big\}\cdot\hat\chi_{\rm 1D}(q_z)}\ . \tag{4.254}$$

Three-dimensional ordering at $\boldsymbol{Q}=0$ sets in at $T=T_{\rm c}$, which occurs when $\hat\chi(\boldsymbol{Q})$ has a pole. The equation for this pole is

$$4\gamma^{-2}J_\perp\,\chi_{\rm 1D}=1\qquad\Rightarrow\qquad\frac{4J_\perp}{k_{\rm B}T_{\rm c}}=\exp(-2J_\parallel/k_{\rm B}T_{\rm c})\ . \tag{4.255}$$

This transcendental equation is equivalent to

$$x\,e^x=\frac{1}{\epsilon} \tag{4.256}$$

where $x=2J_\parallel/k_{\rm B}T_{\rm c}$ and $\epsilon=2J_\perp/J_\parallel$. The solution, for small ϵ, is

$$k_{\rm B}T_{\rm c}=\frac{2J_\parallel}{\ln\big(J_\parallel/2J_\perp\big)}+\ldots\ . \tag{4.257}$$

Thus, $T_{\rm c} > 0$ for all finite $J_\perp$, with $T_{\rm c}$ going to zero rather slowly as $J_\perp \to 0$.

Similar physics is present in the antiferromagnetic insulator phase of the cuprate superconductors. The antiferromagnetic (staggered) susceptibility of the two-dimensional Heisenberg model diverges as $T \to 0$ as $\chi^{\rm stag}_{\rm 2D} \sim J^{-1} \exp(\rho J/k_{\rm B}T)$, where ρ is a dimensionless measure of quantum fluctuations. As in the $d = 1$ Ising case, there is no phase transition at any finite temperature, in this case owing to the HMW theorem. However, when the quasi-2D layers are weakly coupled with antiferromagnetic coupling J' (the base structure is a cubic perovskite), three-dimensional Néel ordering sets in at the antiferromagnetic wavevector $\boldsymbol{Q} = (\pi/a, \pi/a, \pi/c)$ at a critical temperature $T_{\rm N} \approx J/k_{\rm B} \ln(J/J')$.

4.9 Spin Wave Theory

Recall the SU(2) algebra of quantum spin: $[S^\alpha, S^\beta] = i\epsilon_{\alpha\beta\gamma} S^\gamma$ (set $\hbar = 1$ for convenience). Defining $S^\pm = S^x \pm iS^y$, we have, equivalently,

$$[S^z, S^\pm] = \pm\, S^\pm \quad , \quad [S^+, S^-] = 2\, S^z \ . \tag{4.258}$$

The Holstein-Primakoff transformation (1940) maps the spin algebra onto that of a single bosonic oscillator:

$$\begin{aligned} S^+ &= a^\dagger \, (2S - a^\dagger a)^{1/2} \\ S^- &= (2S - a^\dagger a)^{1/2} \, a \\ S^z &= a^\dagger a - S \ . \end{aligned} \tag{4.259}$$

The state $|\, S^z = -S \,\rangle$ is the vacuum $|\, 0 \,\rangle$ in the boson picture. The highest weight state, $|\, S^z = +S \,\rangle$ corresponds to the state $|\, 2S \,\rangle$ in the boson picture, *i.e.* an occupancy of $n = 2S$ bosons.

<u>*EXERCISE:*</u> Verify that the bosonic representation of the spin operators in (4.259) satisfies the SU(2) commutation relations of quantum spin.

What does it mean to take the square root of an operator like $2S - a^\dagger a$? Simple! Just evaluate it in a basis diagonal in $a^\dagger a$, *i.e.* the number basis:

$$a^\dagger a \,|\, n \,\rangle = n \,|\, n \,\rangle \quad \Rightarrow \quad (2S - a^\dagger a)^{1/2} \,|\, n \,\rangle = (2S - n)^{1/2} \,|\, n \,\rangle \ . \tag{4.260}$$

Note that physical boson states are restricted to $n \in \{0, 1, \ldots, 2S\}$. What about states with $n > 2S$? The nice thing here is that we needn't worry about them at all, because S^+, S^-, and S^z do not connect states with $0 \le n \le 2S$ to states with $n > 2S$. For example, when applying the spin raising operator S^+ to the highest weight state $|\, S^z = +S \,\rangle$, in boson language we have

$$S^+ \,\big|\, S^z = +S \,\big\rangle = a^\dagger \, (2S - a^\dagger a)^{1/2} \,\big|\, n = 2S \,\big\rangle = 0 \ , \tag{4.261}$$

as required.

While the HP transformation is exact, it really doesn't buy us anything unless we start making some approximations and derive a systematic expansion in 'spin wave' interactions.

4.9.1 Ferromagnetic Spin Waves

Consider the classical ground state $|\,\mathrm{F}\,\rangle = |\downarrow\downarrow\cdots\downarrow\rangle$ in which all spins are pointing 'down', with $S^z = -S$. In the boson language, the occupancy at each site is zero. This is in fact an eigenstate of the Heisenberg Hamiltonian

$$\mathcal{H} = -\sum_{i<j} J_{ij}\,\boldsymbol{S}_i\cdot\boldsymbol{S}_j \tag{4.262}$$

with eigenvalue $E_0 = -S^2\sum_{i<j}J_{ij}$. If all the interactions are ferromagnetic, *i.e.* $J_{ij} > 0\ \forall\ (i,j)$, then this state clearly is the ground state. We now express the Heisenberg interaction $\boldsymbol{S}_i\cdot\boldsymbol{S}_j$ in terms of the boson creation and annihilation operators. To this end, we perform a Taylor expansion of the radical,

$$(2S - a^\dagger a)^{1/2} = \sqrt{2S}\left\{1 - \frac{1}{2}\left(\frac{a^\dagger a}{2S}\right) - \frac{1}{8}\left(\frac{a^\dagger a}{2S}\right)^2 + \ldots\right\}, \tag{4.263}$$

so that

$$\begin{aligned}
\boldsymbol{S}_i\cdot\boldsymbol{S}_j &= \tfrac{1}{2}\,S_i^+S_j^- + \tfrac{1}{2}\,S_i^-S_j^+ + S_i^zS_j^z \qquad (4.264)\\
&= S\,a_i^\dagger\left(1 - \frac{a_i^\dagger a_i}{4S} + \ldots\right)\left(1 - \frac{a_j^\dagger a_j}{4S} + \ldots\right)a_j \\
&\quad + S\left(1 - \frac{a_i^\dagger a_i}{4S} + \ldots\right)a_i\,a_j^\dagger\left(1 - \frac{a_j^\dagger a_j}{4S} + \ldots\right) + (a_i^\dagger a_i - S)\,(a_j^\dagger a_j - S) \\
&= S^2 + S\Big(a_i^\dagger a_j + a_j^\dagger a_i - a_i^\dagger a_i - a_j^\dagger a_j\Big) + \Big\{a_i^\dagger a_i a_j^\dagger a_j - \tfrac{1}{4}\,a_i^\dagger a_i^\dagger a_i a_j \qquad (4.265)\\
&\qquad\qquad - \tfrac{1}{4}a_i^\dagger a_j^\dagger a_j a_j - \tfrac{1}{4}\,a_j^\dagger a_i^\dagger a_i a_i - \tfrac{1}{4}a_j^\dagger a_j^\dagger a_j a_i\Big\} + \mathcal{O}(1/S)\ .
\end{aligned}$$

Note that a systematic expansion in powers of $1/S$ can be performed. The Heisenberg Hamiltonian now becomes

$$\mathcal{H} = \overbrace{-S^2\sum_{i<j}J_{ij}}^{\substack{\text{classical ground}\\ \text{state energy } \mathcal{O}(S^2)}} + \overbrace{S\sum_{i<j}J_{ij}\left(a_i^\dagger a_i + a_j^\dagger a_j - a_i^\dagger a_j - a_j^\dagger a_i\right)}^{\text{spin-wave Hamiltonian } \mathcal{H}_{\rm sw}} + \overbrace{\mathcal{O}(S^0)}^{\substack{\text{spin-wave}\\ \text{interactions}}}\ . \tag{4.266}$$

We assume our sites are elements of a Bravais lattice, and we Fourier transform:

$$a_i = \frac{1}{\sqrt{N}}\sum_{\boldsymbol{q}} e^{+i\boldsymbol{q}\cdot\boldsymbol{R}_i}\,a_{\boldsymbol{q}} \qquad\qquad a_i^\dagger = \frac{1}{\sqrt{N}}\sum_{\boldsymbol{q}} e^{-i\boldsymbol{q}\cdot\boldsymbol{R}_i}\,a_{\boldsymbol{q}}^\dagger \tag{4.267}$$

$$a_{\boldsymbol{q}} = \frac{1}{\sqrt{N}}\sum_{i} e^{-i\boldsymbol{q}\cdot\boldsymbol{R}_i}\,a_i \qquad\qquad a_{\boldsymbol{q}}^\dagger = \frac{1}{\sqrt{N}}\sum_{i} e^{+i\boldsymbol{q}\cdot\boldsymbol{R}_i}\,a_i^\dagger\ . \tag{4.268}$$

Note that the canonical commutation relations are preserved by this transformation:

$$[a_i, a_j^\dagger] = \delta_{ij} \quad\Longleftrightarrow\quad [a_{\boldsymbol{q}}, a_{\boldsymbol{q}'}^\dagger] = \delta_{\boldsymbol{q}\boldsymbol{q}'}\ . \tag{4.269}$$

Using the result

$$\frac{1}{N}\sum_i e^{i(\boldsymbol{q}-\boldsymbol{q}')\cdot\boldsymbol{R}_i} = \delta_{\boldsymbol{q}\boldsymbol{q}'} \ , \tag{4.270}$$

we obtain the spin-wave Hamiltonian

$$\mathcal{H}_{\rm sw} = S\sum_{\boldsymbol{q}} \left[\hat{J}(0) - \hat{J}(\boldsymbol{q})\right] a^\dagger_{\boldsymbol{q}} a_{\boldsymbol{q}} \ , \tag{4.271}$$

from which we read off the spin-wave dispersion

$$\begin{aligned}\hbar\omega_{\boldsymbol{q}} &= S\big[\hat{J}(0) - \hat{J}(\boldsymbol{q})\big] \\ &= \tfrac{1}{6}S\Big[\sum_{\boldsymbol{R}} J(\boldsymbol{R})\,\boldsymbol{R}^2\Big]\boldsymbol{q}^2 + \mathcal{O}(q^4) \ .\end{aligned} \tag{4.272}$$

The above sum on $\boldsymbol{R}$ converges if $J(R\to\infty) \sim R^{-(d+2+\epsilon)}$ with $\epsilon > 0$.

4.9.2 Static Correlations in the Ferromagnet

The transverse spin-spin correlation function is

$$\begin{aligned}\langle S_i^+ S_j^-\rangle &= \left\langle a_i^\dagger \left(2S - a_i^\dagger a_i\right)^{1/2}\left(2S - a_j^\dagger a_j\right)^{1/2} a_j\right\rangle \\ &= 2S\,\langle a_i^\dagger a_j\rangle + \mathcal{O}(S^0) \\ &= 2S\Omega\int\limits_{\hat\Omega}\frac{d^dk}{(2\pi)^d}\,\frac{e^{i\boldsymbol{k}\cdot(\boldsymbol{R}_j-\boldsymbol{R}_i)}}{e^{\hbar\omega_{\boldsymbol{k}}/k_{\rm B}T}-1} \ .\end{aligned} \tag{4.273}$$

The longitudinal spin-spin correlation function is

$$\langle S_i^z S_j^z\rangle - \langle S_i^z\rangle\langle S_j^z\rangle = \langle a_i^\dagger a_i a_j^\dagger a_j\rangle - \langle a_i^\dagger a_i\rangle\langle a_j^\dagger a_j\rangle = \mathcal{O}(S^0) \ . \tag{4.274}$$

Note that the average spin polarization per site is

$$\begin{aligned}\langle S_i^z\rangle &= -S + \langle a_i^\dagger a_i\rangle \\ &= -S + \Omega\int\limits_{\hat\Omega}\frac{d^dk}{(2\pi)^d}\,\frac{1}{e^{\hbar\omega_{\boldsymbol{k}}/k_{\rm B}T}-1} \ .\end{aligned} \tag{4.275}$$

Now as $k\to 0$ the denominator above vanishes as k^2, hence the average spin polarization per site *diverges* when $d \le 2$. This establishes a "poor man's version" of the HMW theorem: as infinite spin polarization is clearly absurd, there must have been something wrong with our implicit assumption that long-ranged order persists to finite T. In $d = 3$ dimensions, one finds $\langle S_i^z\rangle = -S + \mathcal{O}(T^{3/2})$.

4.9.3 Antiferromagnetic Spin Waves

The case of the ferromagnet is special because the classical ground state $|\,\mathrm{F}\,\rangle$ is in fact a quantum eigenstate – indeed the ground state – of the ferromagnetic Heisenberg Hamiltonian.[11] In the case of the Heisenberg antiferromagnet, this is no longer the case. The ground state itself is a linear combination of classical states. What is the classical ground state? For an antiferromagnet on a bipartite lattice,[12] the classical ground state has each sublattice maximally polarized, with the magnetization on the two sublattices oppositely oriented. Choosing the axis of polarization as $\hat{\boldsymbol{z}}$, this means $S_i^z = -S$ is $i \in A$ and $S_i^z = +S$ if $i \in B$. We'll call this state $|\,\mathrm{N}\,\rangle$, since it is a classical Néel state.

Let is assume that the lattice is a Bravais lattice with a two-element basis described by basis vectors 0 and $\boldsymbol{\delta}$. Thus, if $\boldsymbol{R}$ is any direct lattice vector, an A sublattice site lies at $\boldsymbol{R}$ and a B site at $\boldsymbol{R} + \boldsymbol{\delta}$. The Heisenberg Hamiltonian is written

$$\begin{aligned}\mathcal{H} = -\sum_{\boldsymbol{R},\boldsymbol{R}'} \Big\{ &\tfrac{1}{2}\, J_{\mathrm{AA}}(\boldsymbol{R}-\boldsymbol{R}')\, \boldsymbol{S}_{\mathrm{A}}(\boldsymbol{R}) \cdot \boldsymbol{S}_{\mathrm{A}}(\boldsymbol{R}') + \tfrac{1}{2}\, J_{\mathrm{BB}}(\boldsymbol{R}-\boldsymbol{R}')\, \boldsymbol{S}_{\mathrm{B}}(\boldsymbol{R}) \cdot \boldsymbol{S}_{\mathrm{B}}(\boldsymbol{R}') \\ &+ J_{\mathrm{AB}}(\boldsymbol{R}-\boldsymbol{R}'-\boldsymbol{\delta})\, \boldsymbol{S}_{\mathrm{A}}(\boldsymbol{R}) \cdot \boldsymbol{S}_{\mathrm{B}}(\boldsymbol{R}') \Big\}\ .\end{aligned} \tag{4.276}$$

Here $\boldsymbol{S}_{\mathrm{A}}(\boldsymbol{R})$ represents the spin on the A sublattice located at position $\boldsymbol{R}$, while $\boldsymbol{S}_{\mathrm{B}}(\boldsymbol{R})$ represents the B sublattice spin located at $\boldsymbol{R}+\boldsymbol{\delta}$. The factor of $\frac{1}{2}$ multiplying the J_{AA} and J_{BB} terms avoids double-counting the AA and BB interactions. The Néel state will be the classical ground state if $J_{\mathrm{AA}} > 0$ and $J_{\mathrm{BB}} > 0$ and $J_{\mathrm{AB}} < 0$. It may remain the ground state even if some of the interactions are frustrating, *i.e.* $J_{\mathrm{AA}} < 0$, $J_{\mathrm{BB}} < 0$, and/or $J_{\mathrm{AB}} > 0$ between certain sites.

We'd like the Néel state $|\,\mathrm{N}\,\rangle = |\uparrow\downarrow\uparrow\downarrow\uparrow\ldots\rangle$ to be the vacuum for the Holstein-Primakoff bosons. To accomplish this, we rotate the spin operators on the B sublattice by π about the $\hat{\boldsymbol{y}}$-axis in the internal SU(2) space, sending $S^x \to -S^x$, $S^y \to S^y$, and $S^z \to -S^z$. In the language of HP bosons, we have the following:

A Sublattice	B Sublattice
$S^+ = a^\dagger (2S - a^\dagger a)^{1/2}$	$S^+ = -(2S - b^\dagger b)^{1/2}\, b$
$S^- = (2S - a^\dagger a)^{1/2}\, a$	$S^- = -b^\dagger (2S - b^\dagger b)^{1/2}$
$S^z = a^\dagger a - S$	$S^z = S - b^\dagger b$

(4.277)

[11] Of course, $|\mathrm{F}\rangle$ is also an eigenstate – the highest lying excited state – of the antiferromagnetic Heisenberg Hamiltonian.

[12] A bipartite lattice is one which may be decomposed into two sublattices A and B, such that all the neighbors of any site in A lie in B, and all the neighbors of any site in B lie in A. Examples of bipartite lattices: square, honeycomb, simple cubic, body-centered cubic, hexagonal. Examples of lattices which are not bipartite: triangular, Kagomé, face-centered cubic.

We may now write the Heisenberg interaction as an expansion in powers of $1/S$:

$$\begin{aligned}
\boldsymbol{S}_{\rm A}(\boldsymbol{R})\cdot\boldsymbol{S}_{\rm A}(\boldsymbol{R}') &= S^2 + S\left(a^\dagger_{\boldsymbol{R}}\,a_{\boldsymbol{R}'} + a^\dagger_{\boldsymbol{R}'}\,a_{\boldsymbol{R}} - a^\dagger_{\boldsymbol{R}}\,a_{\boldsymbol{R}} - a^\dagger_{\boldsymbol{R}'}\,a_{\boldsymbol{R}'}\right) + \mathcal{O}(S^0) \\
\boldsymbol{S}_{\rm B}(\boldsymbol{R})\cdot\boldsymbol{S}_{\rm B}(\boldsymbol{R}') &= S^2 + S\left(b^\dagger_{\boldsymbol{R}}\,b_{\boldsymbol{R}'} + b^\dagger_{\boldsymbol{R}'}\,b_{\boldsymbol{R}} - b^\dagger_{\boldsymbol{R}}\,b_{\boldsymbol{R}} - b^\dagger_{\boldsymbol{R}'}\,b_{\boldsymbol{R}'}\right) + \mathcal{O}(S^0) \\
\boldsymbol{S}_{\rm A}(\boldsymbol{R})\cdot\boldsymbol{S}_{\rm B}(\boldsymbol{R}') &= -S^2 + S\left(a^\dagger_{\boldsymbol{R}}\,a_{\boldsymbol{R}} + b^\dagger_{\boldsymbol{R}}\,b_{\boldsymbol{R}} - a^\dagger_{\boldsymbol{R}}\,b^\dagger_{\boldsymbol{R}'} - a_{\boldsymbol{R}}\,b_{\boldsymbol{R}'}\right) + \mathcal{O}(S^0)\ .
\end{aligned} \tag{4.278}$$

Thus, the classical ground state energy is the $\mathcal{O}(S^2)$ term,

$$E_{\rm cl} = S^2\sum_{\boldsymbol{R},\boldsymbol{R}'}\Big\{-\tfrac{1}{2}\,J_{\rm AA}(\boldsymbol{R}-\boldsymbol{R}') - \tfrac{1}{2}\,J_{\rm BB}(\boldsymbol{R}-\boldsymbol{R}') + J_{\rm AB}(\boldsymbol{R}-\boldsymbol{R}'-\boldsymbol{\delta})\Big\}\ . \tag{4.279}$$

The spin-wave Hamiltonian is the $\mathcal{O}(S^1)$ piece,

$$\begin{aligned}
\mathcal{H}_{\rm sw} = -S\sum_{\boldsymbol{R},\boldsymbol{R}'}\Big\{&J_{\rm AA}(\boldsymbol{R}-\boldsymbol{R}')\Big(a^\dagger_{\boldsymbol{R}}\,a_{\boldsymbol{R}'} - a^\dagger_{\boldsymbol{R}}\,a_{\boldsymbol{R}}\Big) + J_{\rm BB}(\boldsymbol{R}-\boldsymbol{R}')\Big(b^\dagger_{\boldsymbol{R}}\,b_{\boldsymbol{R}'} - b^\dagger_{\boldsymbol{R}}\,b_{\boldsymbol{R}}\Big) \\
&+ J_{\rm AB}(\boldsymbol{R}-\boldsymbol{R}'-\boldsymbol{\delta})\Big(a^\dagger_{\boldsymbol{R}}\,a_{\boldsymbol{R}} + b^\dagger_{\boldsymbol{R}}\,b_{\boldsymbol{R}} - a^\dagger_{\boldsymbol{R}}\,b^\dagger_{\boldsymbol{R}'} - a_{\boldsymbol{R}}\,b_{\boldsymbol{R}'}\Big)\Big\}\ .
\end{aligned} \tag{4.280}$$

We now Fourier transform:

$$a_{\boldsymbol{R}} = \frac{1}{\sqrt{N}}\sum_{\boldsymbol{k}} e^{+i\boldsymbol{k}\cdot\boldsymbol{R}}\,a_{\boldsymbol{k}} \qquad\qquad a^\dagger_{\boldsymbol{R}} = \frac{1}{\sqrt{N}}\sum_{\boldsymbol{k}} e^{-i\boldsymbol{k}\cdot\boldsymbol{R}}\,a^\dagger_{\boldsymbol{k}} \tag{4.281}$$

$$b_{\boldsymbol{R}} = \frac{1}{\sqrt{N}}\sum_{\boldsymbol{k}} e^{+i\boldsymbol{k}\cdot(\boldsymbol{R}+\boldsymbol{\delta})}\,b_{\boldsymbol{k}} \qquad\qquad b^\dagger_{\boldsymbol{R}} = \frac{1}{\sqrt{N}}\sum_{\boldsymbol{k}} e^{-i\boldsymbol{k}\cdot(\boldsymbol{R}+\boldsymbol{\delta})}\,b^\dagger_{\boldsymbol{k}}\ , \tag{4.282}$$

which leads to

$$\begin{aligned}
\sum_{\boldsymbol{R},\boldsymbol{R}'} J_{\rm AA}(\boldsymbol{R}-\boldsymbol{R}')\,a^\dagger_{\boldsymbol{R}}\,a_{\boldsymbol{R}'} &= \frac{1}{N}\sum_{\boldsymbol{k},\boldsymbol{k}'}\sum_{\boldsymbol{R},\boldsymbol{R}'} J_{\rm AA}(\boldsymbol{R}-\boldsymbol{R}')e^{i(\boldsymbol{k}'\cdot\boldsymbol{R}'-\boldsymbol{k}\cdot\boldsymbol{R})}\,a^\dagger_{\boldsymbol{k}}a_{\boldsymbol{k}'} \\
&= \sum_{\boldsymbol{k}}\hat{J}_{\rm AA}(\boldsymbol{k})\,a^\dagger_{\boldsymbol{k}}\,a_{\boldsymbol{k}}
\end{aligned} \tag{4.283}$$

$$\begin{aligned}
\sum_{\boldsymbol{R},\boldsymbol{R}'} J_{\rm AB}(\boldsymbol{R}-\boldsymbol{R}'-\boldsymbol{\delta})\,a^\dagger_{\boldsymbol{R}}\,b^\dagger_{\boldsymbol{R}'} &= \frac{1}{N}\sum_{\boldsymbol{k},\boldsymbol{k}'}\sum_{\boldsymbol{R},\boldsymbol{R}'} J_{\rm AB}(\boldsymbol{R}-\boldsymbol{R}'-\boldsymbol{\delta})e^{i\left(\boldsymbol{k}'\cdot(\boldsymbol{R}'+\boldsymbol{\delta})-\boldsymbol{k}\cdot\boldsymbol{R}\right)}\,a^\dagger_{\boldsymbol{k}}b^\dagger_{-\boldsymbol{k}'} \\
&= \sum_{\boldsymbol{k}}\hat{J}_{\rm AB}(\boldsymbol{k})\,a^\dagger_{\boldsymbol{k}}\,b^\dagger_{-\boldsymbol{k}}\ ,
\end{aligned} \tag{4.284}$$

where, assuming $J_{\rm AA}$, $J_{\rm BB}$ and $J_{\rm AB}$ are functions only of the magnitude of their arguments,

$$\begin{aligned}
\hat{J}_{\rm AA}(\boldsymbol{k}) &\equiv \sum_{\boldsymbol{R}} J_{\rm AA}\big(|\boldsymbol{R}|\big)\,e^{i\boldsymbol{k}\cdot\boldsymbol{R}} \\
\hat{J}_{\rm BB}(\boldsymbol{k}) &\equiv \sum_{\boldsymbol{R}} J_{\rm BB}\big(|\boldsymbol{R}|\big)\,e^{i\boldsymbol{k}\cdot\boldsymbol{R}} \\
\hat{J}_{\rm AB}(\boldsymbol{k}) &\equiv \sum_{\boldsymbol{R}} J_{\rm AB}\big(|\boldsymbol{R}+\delta|\big)\,e^{i\boldsymbol{k}\cdot(\boldsymbol{R}+\boldsymbol{\delta})}\ .
\end{aligned} \tag{4.285}$$

Note that $\hat{J}_{\rm AA}(\boldsymbol{k}) = \hat{J}_{\rm AA}(-\boldsymbol{k}) = \big[\hat{J}_{\rm AA}(\boldsymbol{k})\big]^*$ (similarly for $J_{\rm BB}$), and $\hat{J}_{\rm AB}(\boldsymbol{k}) = \big[\hat{J}_{\rm AB}(-\boldsymbol{k})\big]^*$.

The spin-wave Hamiltonian may now be written as

$$\mathcal{H}_{\rm sw} = S\sum_{\boldsymbol{k}}\Bigg\{\Big(\hat{J}_{\rm AA}(0) - \hat{J}_{\rm AA}(\boldsymbol{k}) - \hat{J}_{\rm AB}(0)\Big)\,a^\dagger_{\boldsymbol{k}}\,a_{\boldsymbol{k}} + \Big(\hat{J}_{\rm BB}(0) - \hat{J}_{\rm BB}(\boldsymbol{k}) - \hat{J}_{\rm AB}(0)\Big)\,b^\dagger_{\boldsymbol{k}}\,b_{\boldsymbol{k}} + \hat{J}_{\rm AB}(\boldsymbol{k})\,a^\dagger_{\boldsymbol{k}}\,b^\dagger_{-\boldsymbol{k}} + J^*_{\rm AB}(\boldsymbol{k})\,a_{\boldsymbol{k}}\,b_{-\boldsymbol{k}}\Bigg\}\ . \tag{4.286}$$

In other words,

$$\boxed{\mathcal{H}_{\rm sw} = \sum_{\boldsymbol{k}}\Big\{\Omega^{\rm AA}_{\boldsymbol{k}}\,a^\dagger_{\boldsymbol{k}}\,a_{\boldsymbol{k}} + \Omega^{\rm BB}_{\boldsymbol{k}}\,b^\dagger_{\boldsymbol{k}}\,b_{\boldsymbol{k}} + \Delta_{\boldsymbol{k}}\,a^\dagger_{\boldsymbol{k}}\,b^\dagger_{-\boldsymbol{k}} + \Delta^*_{\boldsymbol{k}}\,a_{\boldsymbol{k}}\,b_{-\boldsymbol{k}}\Big\}} \tag{4.287}$$

with

$$\begin{aligned}\Omega^{\rm AA}_{\boldsymbol{k}} &= S\Big(\hat{J}_{\rm AA}(0) - \hat{J}_{\rm AA}(\boldsymbol{k}) - \hat{J}_{\rm AB}(0)\Big)\\ \Omega^{\rm BB}_{\boldsymbol{k}} &= S\Big(\hat{J}_{\rm BB}(0) - \hat{J}_{\rm BB}(\boldsymbol{k}) - \hat{J}_{\rm AB}(0)\Big)\end{aligned} \tag{4.288}$$

and

$$\Delta_{\boldsymbol{k}} = S\,\hat{J}_{\rm AB}(\boldsymbol{k})\ . \tag{4.289}$$

Henceforth we shall assume $J_{\rm AA}(\boldsymbol{R}) = J_{\rm BB}(\boldsymbol{R})$, so $\Omega^{\rm AA}_{\boldsymbol{k}} = \Omega^{\rm BB}_{\boldsymbol{k}} \equiv \Omega_{\boldsymbol{k}}$.

Note that the vacuum $\big|\,0\,\big\rangle$ for the a and b bosons is not an eigenstate of $\mathcal{H}_{\rm sw}$, owing to the spin-wave pair creation term $\Delta^*_{\boldsymbol{k}}\,a_{\boldsymbol{k}}\,b_{-\boldsymbol{k}}$. This can be traced back to the effect on the Néel state of the Heisenberg interaction,

$$\boldsymbol{S}_i\cdot\boldsymbol{S}_j = \tfrac{1}{2}\,S^+_i\,S^-_j + \tfrac{1}{2}\,S^-_i\,S^+_j + S^z_i\,S^z_j\ . \tag{4.290}$$

If $i\in{\rm A}$ and $j\in{\rm B}$, then the term $S^+_i\,S^-_j$ acts on the configuration $|\,-S\,,+S\,\rangle$ and converts it to $2S\,|\,-S+1\,,\,S-1\,\rangle$. Nevertheless, we can diagonalize $\mathcal{H}_{\rm sw}$ by means of a canonical (but not unitary!) transformation, known as the *Bogoliubov transformation*. Note that for each $\boldsymbol{k}\in\hat{\Omega}$, the spin-wave Hamiltonian couples only four operators: $a^\dagger_{\boldsymbol{k}}$, $a_{\boldsymbol{k}}$, $b^\dagger_{-\boldsymbol{k}}$, and $b_{-\boldsymbol{k}}$. We write the Bogoliubov transformation as

$$a_{\boldsymbol{k}} = u_{\boldsymbol{k}}\,\alpha_{\boldsymbol{k}} - v^*_{\boldsymbol{k}}\,\beta^\dagger_{-\boldsymbol{k}} \qquad\qquad b_{-\boldsymbol{k}} = u_{\boldsymbol{k}}\,\beta_{-\boldsymbol{k}} - v^*_{\boldsymbol{k}}\,\alpha^\dagger_{\boldsymbol{k}} \tag{4.291}$$

$$a^\dagger_{\boldsymbol{k}} = u^*_{\boldsymbol{k}}\,\alpha^\dagger_{\boldsymbol{k}} - v_{\boldsymbol{k}}\,\beta_{-\boldsymbol{k}} \qquad\qquad b^\dagger_{-\boldsymbol{k}} = u^*_{\boldsymbol{k}}\,\beta^\dagger_{-\boldsymbol{k}} - v_{\boldsymbol{k}}\,\alpha_{\boldsymbol{k}} \tag{4.292}$$

One can readily verify that this transformation preserves the canonical bosonic commutation relations,

$$\big[a_{\boldsymbol{k}}, a^\dagger_{\boldsymbol{k}'}\big] = \big[b_{\boldsymbol{k}}, b^\dagger_{\boldsymbol{k}'}\big] = \big[\alpha_{\boldsymbol{k}}, \alpha^\dagger_{\boldsymbol{k}'}\big] = \big[\beta_{\boldsymbol{k}}, \beta^\dagger_{\boldsymbol{k}'}\big] = \delta_{\boldsymbol{k}\boldsymbol{k}'} \tag{4.293}$$

provided that

$$u^*_{\boldsymbol{k}}\,u_{\boldsymbol{k}} - v^*_{\boldsymbol{k}}\,v_{\boldsymbol{k}} = 1\ . \tag{4.294}$$

The inverse transformation is

$$\alpha_{\boldsymbol{k}} = u_{\boldsymbol{k}}^* \, a_{\boldsymbol{k}} + v_{\boldsymbol{k}}^* \, b_{-\boldsymbol{k}}^\dagger \qquad\qquad \beta_{-\boldsymbol{k}} = u_{\boldsymbol{k}}^* \, b_{-\boldsymbol{k}} + v_{\boldsymbol{k}}^* \, a_{\boldsymbol{k}}^\dagger \tag{4.295}$$

$$\alpha_{\boldsymbol{k}}^\dagger = u_{\boldsymbol{k}} \, a_{\boldsymbol{k}}^\dagger + v_{\boldsymbol{k}} \, b_{-\boldsymbol{k}} \qquad\qquad \beta_{-\boldsymbol{k}}^\dagger = u_{\boldsymbol{k}} \, b_{-\boldsymbol{k}}^\dagger + v_{\boldsymbol{k}} \, a_{\boldsymbol{k}} \ . \tag{4.296}$$

We'll write

$$u_{\boldsymbol{k}} = \exp(i\eta_{\boldsymbol{k}}) \, \cosh(\theta_{\boldsymbol{k}}) \qquad , \qquad v_{\boldsymbol{k}} = \exp(-i\eta_{\boldsymbol{k}}) \, \sinh(\theta_{\boldsymbol{k}}) \ . \tag{4.297}$$

We may then write

$$\begin{aligned} a_{\boldsymbol{k}} &= \exp(i\eta_{\boldsymbol{k}}) \, \cosh(\theta_{\boldsymbol{k}}) \, \alpha_{\boldsymbol{k}} - \exp(i\eta_{\boldsymbol{k}}) \, \sinh(\theta_{\boldsymbol{k}}) \, \beta_{-\boldsymbol{k}}^\dagger \\ b_{-\boldsymbol{k}} &= \exp(i\eta_{\boldsymbol{k}}) \, \cosh(\theta_{\boldsymbol{k}}) \, \beta_{-\boldsymbol{k}} - \exp(i\eta_{\boldsymbol{k}}) \, \sinh(\theta_{\boldsymbol{k}}) \, \alpha_{\boldsymbol{k}}^\dagger \end{aligned} \tag{4.298}$$

as well as the inverse

$$\begin{aligned} \alpha_{\boldsymbol{k}} &= \exp(-i\eta_{\boldsymbol{k}}) \, \cosh(\theta_{\boldsymbol{k}}) \, a_{\boldsymbol{k}} + \exp(i\eta_{\boldsymbol{k}}) \, \sinh(\theta_{\boldsymbol{k}}) \, \beta_{-\boldsymbol{k}}^\dagger \\ \beta_{-\boldsymbol{k}} &= \exp(-i\eta_{\boldsymbol{k}}) \, \cosh(\theta_{\boldsymbol{k}}) \, \beta_{-\boldsymbol{k}} + \exp(i\eta_{\boldsymbol{k}}) \, \sinh(\theta_{\boldsymbol{k}}) \, a_{\boldsymbol{k}}^\dagger \ . \end{aligned} \tag{4.299}$$

Substituting into the expressions from $\mathcal{H}_{\rm sw}$, we find

$$\begin{aligned} \Omega_{\boldsymbol{k}} \left(a_{\boldsymbol{k}}^\dagger \, a_{\boldsymbol{k}} + b_{\boldsymbol{k}}^\dagger \, b_{\boldsymbol{k}} \right) = \Omega_{\boldsymbol{k}} \, \cosh(2\theta_{\boldsymbol{k}}) \left(\alpha_{\boldsymbol{k}}^\dagger \, \alpha_{\boldsymbol{k}} + \beta_{-\boldsymbol{k}}^\dagger \, \beta_{-\boldsymbol{k}} + 1 \right) - \Omega_{\boldsymbol{k}} \\ - \Omega_{\boldsymbol{k}} \, \sinh(2\theta_{\boldsymbol{k}}) \left(\alpha_{\boldsymbol{k}}^\dagger \, \beta_{-\boldsymbol{k}}^\dagger + \alpha_{\boldsymbol{k}} \, \beta_{-\boldsymbol{k}} \right) \end{aligned} \tag{4.300}$$

and

$$\begin{aligned} \Delta_{\boldsymbol{k}} \, a_{\boldsymbol{k}}^\dagger \, a_{-\boldsymbol{k}}^\dagger + \Delta_{\boldsymbol{k}}^* \, a_{\boldsymbol{k}} \, b_{-\boldsymbol{k}} = -\left|\Delta_{\boldsymbol{k}}\right| \, \sinh(2\theta_{\boldsymbol{k}}) \left(\alpha_{\boldsymbol{k}}^\dagger \, \alpha_{\boldsymbol{k}} + \beta_{-\boldsymbol{k}}^\dagger \, \beta_{-\boldsymbol{k}} + 1 \right) \\ + \left|\Delta_{\boldsymbol{k}}\right| \, \cosh(2\theta_{\boldsymbol{k}}) \left(\alpha_{\boldsymbol{k}}^\dagger \, \beta_{-\boldsymbol{k}}^\dagger + \alpha_{\boldsymbol{k}} \, \beta_{-\boldsymbol{k}} \right) \ , \end{aligned} \tag{4.301}$$

where we have taken $\eta_{\boldsymbol{k}} = \frac{1}{2} \arg(\Delta_{\boldsymbol{k}})$. Up until now, $\theta_{\boldsymbol{k}}$ has been arbitrary. We now use this freedom to specify $\theta_{\boldsymbol{k}}$ such that the $\left(\alpha_{\boldsymbol{k}}^\dagger \, \beta_{-\boldsymbol{k}}^\dagger + \alpha_{\boldsymbol{k}} \, \beta_{-\boldsymbol{k}} \right)$ terms vanish from $\mathcal{H}_{\rm sw}$. This requires

$$\left|\Delta_{\boldsymbol{k}}\right| \, \cosh(2\theta_{\boldsymbol{k}}) - \Omega_{\boldsymbol{k}} \, \sinh(2\theta_{\boldsymbol{k}}) = 0 \quad \Longrightarrow \quad \tanh(2\theta_{\boldsymbol{k}}) = \frac{\left|\Delta_{\boldsymbol{k}}\right|}{\Omega_{\boldsymbol{k}}} \ , \tag{4.302}$$

which means

$$\cosh(2\theta_{\boldsymbol{k}}) = \frac{\Omega_{\boldsymbol{k}}}{E_{\boldsymbol{k}}} \quad , \quad \sinh(2\theta_{\boldsymbol{k}}) = \frac{\left|\Delta_{\boldsymbol{k}}\right|}{E_{\boldsymbol{k}}} \tag{4.303}$$

along with the dispersion relation

$$E_{\boldsymbol{k}} = \sqrt{\Omega_{\boldsymbol{k}}^2 - \left|\Delta_{\boldsymbol{k}}\right|^2} \ . \tag{4.304}$$

Finally, we may write the diagonalized spin-wave Hamiltonian as

$$\boxed{\mathcal{H}_{\rm sw} = \sum_{\boldsymbol{k}} E_{\boldsymbol{k}} \left(\alpha_{\boldsymbol{k}}^\dagger \, \alpha_{\boldsymbol{k}} + \beta_{\boldsymbol{k}}^\dagger \, \beta_{\boldsymbol{k}} \right) + \sum_{\boldsymbol{k}} \left(E_{\boldsymbol{k}} - \Omega_{\boldsymbol{k}} \right) \ .} \tag{4.305}$$

Note that $E_{\boldsymbol{k}} = E_{-\boldsymbol{k}}$ since $\hat{J}_{\rm AB}(\boldsymbol{k}) = \hat{J}^*_{\rm AB}(-\boldsymbol{k})$. The two terms above represent, respectively, the spin-wave excitation Hamiltonian, and the $\mathcal{O}(S^1)$ quantum correction to the ground state energy. Since $E_{\boldsymbol{k}} < \Omega_{\boldsymbol{k}}$, this correction is always negative.

As $\boldsymbol{k} \to 0$, we have, assuming cubic or higher symmetry,

$$\begin{aligned}\Omega_{\boldsymbol{k}} &= -S \sum_{\boldsymbol{R}} J_{\rm AB}\big(|\boldsymbol{R}+\boldsymbol{\delta}|\big) + \tfrac{1}{6}\, S\, \boldsymbol{k}^2 \sum_{\boldsymbol{R}} J_{\rm AA}\big(|\boldsymbol{R}|\big)\, \boldsymbol{R}^2 + \dots \\ &\equiv SW + SX\boldsymbol{k}^2 + \dots\end{aligned} \tag{4.306}$$

and

$$\begin{aligned}\Delta_{\boldsymbol{k}} &= +S \sum_{\boldsymbol{R}} J_{\rm AB}\big(|\boldsymbol{R}+\boldsymbol{\delta}|\big) - \tfrac{1}{6}\, S\, \boldsymbol{k}^2 \sum_{\boldsymbol{R}} J_{\rm AB}\big(|\boldsymbol{R}+\boldsymbol{\delta}|\big)\, \big|\boldsymbol{R}+\boldsymbol{\delta}\big|^2 + \dots \\ &\equiv -SW + SY\boldsymbol{k}^2 + \dots\ .\end{aligned} \tag{4.307}$$

The energy dispersion is linear: $E_{\boldsymbol{k}} = \hbar c|\boldsymbol{k}|$, where $c = S\sqrt{2W(X+Y)}$. Antiferromagnetic spin waves are Goldstone bosons corresponding to the broken continuous symmetry of global spin rotation. The dispersion vanishes linearly as $\boldsymbol{k} \to 0$, in contrast to the case of ferromagnetic spin waves, where $E_{\boldsymbol{k}}$ vanishes quadratically.

Reduction in Sublattice Magnetization

Let's compute the average of S^z for a spin on the A sublattice:

$$\begin{aligned}\langle S^z(\boldsymbol{R})\rangle &= -S + \langle a^\dagger_{\boldsymbol{R}}\, a_{\boldsymbol{R}}\rangle \\ &= -S + \frac{1}{N}\sum_{\boldsymbol{k}} \langle a^\dagger_{\boldsymbol{k}}\, a_{\boldsymbol{k}}\rangle \\ &= -S + \frac{1}{N}\sum_{\boldsymbol{k}} \big\langle (u^*_{\boldsymbol{k}}\,\alpha^\dagger_{\boldsymbol{k}} - v_{\boldsymbol{k}}\,\beta_{-\boldsymbol{k}})\,(u_{\boldsymbol{k}}\,\alpha_{\boldsymbol{k}} - v^*_{\boldsymbol{k}}\,\beta^\dagger_{-\boldsymbol{k}})\big\rangle \\ &= -S + v_0 \int_{\rm BZ} \frac{d^d k}{(2\pi)^d} \left\{ \frac{\Omega_{\boldsymbol{k}}}{E_{\boldsymbol{k}}}\, \frac{1}{\exp(E_{\boldsymbol{k}}/k_{\rm B}T) - 1} + \frac{1}{2}\left(\frac{\Omega_{\boldsymbol{k}}}{E_{\boldsymbol{k}}} - 1\right)\right\}\ ,\end{aligned} \tag{4.308}$$

where v_0 is the Wigner-Seitz cell volume, and the integral is over the first Brillouin zone. The deviation $\delta S^z = \langle a^\dagger\, a\rangle$ from the classical value $\langle S^z\rangle = -S$ is due to thermal and quantum fluctuations. Note that even at $T = 0$, when the thermal fluctuations vanish, there is still a reduction in sublattice magnetization due to quantum fluctuations. The Néel state satisfies the $S^z_i S^z_j$ part of the Heisenberg interaction, but the full interaction prefers neighboring spins to be arranged in *singlets*, which involves fluctuations about local Néel order.

We've seen that $\Omega_{\boldsymbol{k}} \simeq SW$ and $E_{\boldsymbol{k}} \simeq \hbar c\,|\boldsymbol{k}|$ as $\boldsymbol{k} \to 0$. Thus, the integrand behaves as $T/\boldsymbol{k}^2$ for the first term and as $1/|\boldsymbol{k}|$ for the second term. The integral therefore diverges in $d \le 2$ at finite T and in $d = 1$ even at $T = 0$. Thermal and quantum fluctuations melt the classical ordered state.

4.9.4 Specific Heat due to Spin Waves

The long wavelength dispersion $\omega_{\boldsymbol{q}} = Aq^2$ has thermodynamic consequences. Consider a general case of a bosonic dispersion $\omega_{\boldsymbol{q}} = A|\boldsymbol{q}|^\sigma$. The internal energy for a system in d space dimensions is then

$$\begin{aligned} E(T) &= V \int \frac{d^d k}{(2\pi)^d} \, \frac{Ak^\sigma}{e^{\beta A q^\sigma} - 1} \\ &= \frac{AV\Omega_d}{(2\pi)^d} \Big(\frac{k_{\rm B} T}{A} \Big)^{1+\frac{d}{\sigma}} \int\limits_0^\infty du \, \frac{u^{d/\sigma}}{e^u - 1} \end{aligned} \tag{4.309}$$

where $\Omega_d = 2\pi^{d/2}/\Gamma(d/2)$ is the area of the unit sphere in d dimensions. Thus, $E(T) \propto T^{1+\frac{d}{\sigma}}$, leading to a low-temperature heat capacity of

$$C_V = \Gamma(2 + \tfrac{1}{2}d) \, \zeta(1 + \tfrac{1}{2}d) \, \frac{k_{\rm B} V \Omega_d}{(2\pi)^d} \left(\frac{k_{\rm B} T}{A} \right)^{d/\sigma} . \tag{4.310}$$

At high T, one must impose a cutoff at the edge of the Brillouin zone, where $k \sim \pi/a$, in order not to overcount the modes. One finds

$$E(T) = k_{\rm B} T \, V \int\limits_{\hat{\Omega}} \frac{d^d k}{(2\pi)^d} = N k_{\rm B} T \ , \tag{4.311}$$

where N is the number of unit cells. This simply is the Dulong-Petit result of $k_{\rm B}T$ per mode.

For ferromagnetic spin waves, we found $\sigma = 2$, hence $C_V \propto T^{d/2}$ at low temperatures. As we shall see, for antiferromagnetic spin waves, one has $\sigma = 1$, as in the case of acoustic phonons, hence $C_V \propto T^d$.

Suppose we write the long-wavelength ferromagnetic spin-wave dispersion as $\hbar\omega_{\boldsymbol{q}} = CJ(\boldsymbol{q}a)^2$, where a is the lattice spacing, J is the nearest neighbor exchange, and C is a dimensionless constant. The ferromagnetic low-temperature specific heat is then

$$C_V^{\rm F} = \Gamma(2 + \tfrac{1}{2}d) \, \zeta(1 + \tfrac{1}{2}d) \, \frac{k_{\rm B} V \Omega_d}{(2\pi a)^d} \left(\frac{k_{\rm B} T}{CJ} \right)^{d/2} , \tag{4.312}$$

hence $C_V^{\rm F} \propto (T/\Theta_J)^{d/2}$, with $\Theta_J \equiv CJ/k_{\rm B}$. Acoustic phonons with a $\omega_{\boldsymbol{k}} = \hbar c|\boldsymbol{k}|$ dispersion lead to a Debye heat capacity

$$C_V^{\rm D} = \Gamma(2 + d) \, \zeta(1 + d) \, \frac{k_{\rm B} V \Omega_d}{(2\pi a)^d} \left(\frac{k_{\rm B} T}{\hbar c/a} \right)^{d} , \tag{4.313}$$

hence $C^{\rm D} \propto (T/\Theta_{\rm D})^d$, with $\Theta_{\rm D} \equiv \hbar c/a k_{\rm B}$. Thus, at the lowest temperatures, the specific heat due to spin waves dominates, but at intermediate temperatures it is the phonon specific heat which dominates. The temperature scale T^* at which the two contributions are roughly equal is given by

$$(T^*/\Theta_J)^{d/2} \simeq (T^*/\Theta_{\rm D})^d \quad \Longrightarrow \quad T^* \simeq \Theta_{\rm D}^2/\Theta_J \ . \tag{4.314}$$

4.10 Appendix: The Foldy-Wouthuysen Transformation

Let us write

$$\mathcal{H} = mc^2\,\gamma^0 + c\gamma^0\,\boldsymbol{\gamma}\cdot\boldsymbol{\pi} + V\ , \tag{4.315}$$

where

$$\boldsymbol{\pi} = \boldsymbol{p} + \tfrac{e}{c}\boldsymbol{A} \tag{4.316}$$

is the dynamical momentum and where the γ^μ are the Dirac matrices,

$$\gamma^0 = \begin{pmatrix} 1_{2\times2} & 0_{2\times2} \\ 0_{2\times2} & -1_{2\times2} \end{pmatrix} \qquad , \qquad \boldsymbol{\gamma} = \begin{pmatrix} 0_{2\times2} & \boldsymbol{\sigma}_{2\times2} \\ -\boldsymbol{\sigma}_{2\times2} & 0_{2\times2} \end{pmatrix}\ . \tag{4.317}$$

Here $\boldsymbol{\sigma}$ is the vector of Pauli matrices.

The idea behind the FW transformation is to unitarily transform to a different Hilbert space basis such that the coupling in $\mathcal{H}$ between the upper and lower components of the Dirac spinor vanishes. This may be done systematically as an expansion in inverse powers of the electron mass m. We begin by defining $K \equiv c\gamma^0\boldsymbol{\gamma}\cdot\boldsymbol{\pi} + V$ so that $\mathcal{H} = mc^2\,\gamma^0 + K$. Note that K is of order m^0. We then write

$$\begin{aligned}\tilde{\mathcal{H}} &= e^{iS}\,\mathcal{H}\,e^{-iS} \\ &= \mathcal{H} + i\big[S,\mathcal{H}\big] + \frac{(i)^2}{2!}\,\big[S,[S,\mathcal{H}]\big] + \dots\ ,\end{aligned} \tag{4.318}$$

where S itself is written as a power series in $(mc^2)^{-1}$:

$$S = \frac{S_0}{mc^2} + \frac{S_1}{(mc^2)^2} + \dots\ . \tag{4.319}$$

The job now is to write $\tilde{\mathcal{H}}$ as a power series in m^{-1}. The first few terms are easy to find:

$$\tilde{\mathcal{H}} = mc^2\,\gamma^0 + K + i\big[S_0,\gamma^0\big] + \frac{1}{mc^2}\left(\big[S_0,K\big] + \big[S_1,\gamma^0\big] - \tfrac{1}{2}\big[S_0,[S_0,K]\big]\right) + \dots \tag{4.320}$$

We choose the operators S_n so as to cancel, at each order in m^{-1}, the off-diagonal terms in $\tilde{\mathcal{H}}$ that couple the upper two components of Ψ to the lower two components of Ψ. To order m^0, we then demand

$$c\gamma^0\boldsymbol{\gamma}\cdot\boldsymbol{\pi} + i\big[S_0,\gamma^0\big] = 0\ . \tag{4.321}$$

Note that we do not demand that $i\big[S_0,\gamma^0\big]$ completely cancel K – indeed it is impossible to find such an S_0, and one way to see this is to take the trace. The trace of any commutator must vanish, but $\mathrm{Tr}\,K = 4V$, which is in general nonzero. But this is of no concern to us, since we only need cancel the (traceless) *off-diagonal* part of K, which is to say $c\gamma^0\boldsymbol{\gamma}\cdot\boldsymbol{\pi}$.

To solve for S_0, one can write it in terms of its four 2×2 subblocks, compute the commutator with γ^0, and then impose eqn. 4.321. One then finds $S_0 = -\frac{i}{2}c\,\boldsymbol{\gamma}\cdot\boldsymbol{\pi}$.

STUDENT EXERCISE: Derive the result $S_0 = -\frac{i}{2}c\,\boldsymbol{\gamma}\cdot\boldsymbol{\pi}$.

At the next level, we have to deal with the term in the round brackets in eqn. 4.320. Since we know S_0, we can compute the first and the third terms therein. In general, this will leave us with an off-diagonal term coupling upper and lower components of Ψ. We then choose S_1 so as to cancel this term. This calculation already is tedious, and we haven't even gotten to the spin-orbit interaction term yet, since it is of order m^{-2} – yecch!

4.10.1 Derivation of the Spin-Orbit Interaction

Here's a simpler way to proceed to order m^{-2}. Let a, b be block indices and i, j be indices within each block. Thus, the component Ψ_{ai} is the $i^{\rm th}$ component of the $a^{\rm th}$ block; $\Psi_{a=1,i=2}$ is the lower component of the upper block, *i.e.* the second component of the four-vector Ψ.

Write the Hamiltonian as

$$\mathcal{H} = mc^2\,\tau^z + c\boldsymbol{\sigma}\cdot\boldsymbol{\pi}\,\tau^x + V(\boldsymbol{r})\ , \tag{4.322}$$

where τ^μ are Pauli matrices with indices a, b and σ^ν are Pauli matrices with indices i, j. The σ and τ matrices commute because they act on different indices.

A very important result regarding Pauli matrices:

$$e^{i\theta\,\hat{\boldsymbol{n}}\cdot\boldsymbol{\tau}/2}\,\tau^\alpha\,e^{-i\theta\,\hat{\boldsymbol{n}}\cdot\boldsymbol{\tau}/2} = n^\alpha n^\beta\,\tau^\beta + \cos\theta\,(\delta^{\alpha\beta} - n^\alpha n^\beta)\,\tau^\beta + \sin\theta\,\epsilon^{\alpha\beta\gamma}\,n^\beta\,\tau^\gamma\ . \tag{4.323}$$

STUDENT EXERCISE: Verify and interpret the above result.

Using this result, we can write

$$A\,\tau^z + B\,\tau^x = \sqrt{A^2+B^2}\cdot e^{-i\tan^{-1}(B/A)\,\tau^y/2}\,\tau^z\,e^{i\tan^{-1}(B/A)\,\tau^y/2}\ , \tag{4.324}$$

and, for our specific purposes,

$$mc^2\,\tau^z + \boldsymbol{\sigma}\cdot\boldsymbol{\pi}\,\tau^x = \sqrt{(mc^2)^2 + (c\boldsymbol{\sigma}\cdot\boldsymbol{\pi})^2}\cdot U\,\tau^z\,U^\dagger\ , \tag{4.325}$$

where

$$U = e^{-i\tan^{-1}(\frac{\boldsymbol{\sigma}\cdot\boldsymbol{\pi}}{mc})\,\tau^y/2}\ . \tag{4.326}$$

The fact that $\boldsymbol{\sigma}\cdot\boldsymbol{\pi}$ is an operator is no obstacle here, since it commutes with the τ^μ matrices. We can give meaning to expressions like $\tan^{-1}(\boldsymbol{\sigma}\cdot\boldsymbol{\pi}/mc)$ in terms of their Taylor series expansions.

We therefore have the result,

$$U^\dagger\,\mathcal{H}\,U = \sqrt{(mc^2)^2 + (c\boldsymbol{\sigma}\cdot\boldsymbol{\pi})^2}\cdot\tau^z + U^\dagger\,V(\boldsymbol{r})\,U\ . \tag{4.327}$$

The first term is diagonal in the block indices. Expanding the square root, we have

$$\begin{aligned} mc^2\sqrt{1+\Big(\frac{\boldsymbol{\sigma}\cdot\boldsymbol{\pi}}{mc}\Big)^2} &= mc^2 + \frac{(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})^2}{2m} + \mathcal{O}(m^{-3}) \\ &= mc^2 + \frac{\boldsymbol{\pi}^2}{2m} + \frac{e\hbar}{2mc}\,\boldsymbol{B}\cdot\boldsymbol{\sigma} + \mathcal{O}(m^{-3})\ , \end{aligned} \tag{4.328}$$

since

$$\begin{aligned}(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})^2 &= \sigma^\mu\sigma^\nu\,\pi^\mu\pi^\nu \\ &= (\delta^{\mu\nu} + i\epsilon^{\mu\nu\lambda}\sigma^\lambda)\,\pi^\mu\pi^\nu \\ &= \boldsymbol{\pi}^2 + \tfrac{i}{2}\epsilon^{\mu\nu\lambda}\left[p^\mu + \tfrac{e}{c}A^\mu, p^\nu + \tfrac{e}{c}A^\nu\right] \\ &= \boldsymbol{\pi}^2 + \frac{e\hbar}{c}\,\boldsymbol{B}\cdot\boldsymbol{\sigma}\ .\end{aligned} \tag{4.329}$$

We next need to compute $U^\dagger\,V(\boldsymbol{r})\,U$ to order m^{-2}. To do this, first note that

$$U = 1 - \frac{i}{2}\,\frac{\boldsymbol{\sigma}\cdot\boldsymbol{\pi}}{mc}\,\tau^y - \frac{1}{8}\Big(\frac{\boldsymbol{\sigma}\cdot\boldsymbol{\pi}}{mc}\Big)^2 + \dots\ , \tag{4.330}$$

Thus,

$$U^\dagger\,V\,U = V + \frac{i}{2mc}\big[\boldsymbol{\sigma}\cdot\boldsymbol{\pi}, V\big]\,\tau^y - \frac{1}{8m^2c^2}\big[\boldsymbol{\sigma}\cdot\boldsymbol{\pi}, [\boldsymbol{\sigma}\cdot\boldsymbol{\pi}, V]\big] + \dots\ . \tag{4.331}$$

Upon reflection, one realizes that, to this order, it suffices to take the first term in the Taylor expansion of $\tan^{-1}(\boldsymbol{\sigma}\cdot\boldsymbol{\pi}/mc)$ in eqn. 4.326, in which case one can then invoke eqn. 4.318 to obtain the above result. The second term on the RHS of eqn. 4.331 is simply $\frac{\hbar}{2mc}\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,V\,\tau^y$. The third term is

$$\begin{aligned}\frac{i\hbar}{8m^2c^2}\big[\sigma^\mu\pi^\mu, \sigma^\nu\partial^\nu V\big] &= \frac{i\hbar}{8m^2c^2}\Big\{\sigma^\mu\left[\pi^\mu, \sigma^\nu\partial^\nu V\right] + \big[\sigma^\mu, \sigma^\nu\partial^\nu V\big]\,\pi^\mu\Big\} \\ &= \frac{i\hbar}{8m^2c^2}\left\{\frac{\hbar}{i}\partial^\mu\partial^\nu V\,\sigma^\mu\sigma^\nu + 2i\epsilon^{\mu\nu\lambda}\sigma^\lambda\partial^\nu V\,\pi^\mu\right\} \\ &= \frac{\hbar^2}{8m^2c^2}\,\boldsymbol{\nabla}^2 V + \frac{\hbar}{4m^2c^2}\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,V\times\boldsymbol{\pi}\ .\end{aligned} \tag{4.332}$$

Therefore,

$$\begin{aligned}U^\dagger\,\mathcal{H}\,U = \left(mc^2 + \frac{\boldsymbol{\pi}^2}{2m} + \frac{e\hbar}{2mc}\,\boldsymbol{B}\cdot\boldsymbol{\sigma}\right)\tau^z + V + \frac{\hbar}{2mc}\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,V\,\tau^y \\ + \frac{\hbar^2}{8m^2c^2}\,\boldsymbol{\nabla}^2 V + \frac{\hbar}{4m^2c^2}\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,V\times\boldsymbol{\pi} + \mathcal{O}(m^{-3})\ .\end{aligned} \tag{4.333}$$

This is not block-diagonal, owing to the last term on the RHS of the top line. We can eliminate this term by effecting yet another unitary transformation. However, this will result in a contribution to the energy of order m^{-3}, so we can neglect it. To substantiate this last claim, drop all the block-diagonal terms except for the leading order one, $mc^2\,\tau^z$, and consider the Hamiltonian

$$\mathcal{K} = mc^2\,\tau^z + \frac{\hbar}{2mc}\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,V\,\tau^y\ . \tag{4.334}$$

We now know how to bring this to block-diagonal form. The result is

$$\begin{aligned}\tilde{\mathcal{K}} &= mc^2\sqrt{1 + \left(\frac{\hbar\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,V}{2m^2c^3}\right)^2}\,\tau^z \\ &= \left(mc^2 + \frac{\hbar^2(\boldsymbol{\nabla}\,V)^2}{8m^3c^4} + \dots\right)\tau^z\ ,\end{aligned} \tag{4.335}$$

and the correction is of order m^{-3}, as promised.

We now assume all the negative energy ($\tau^z = -1$) states are filled. The Hamiltonian for the electrons, valid to $\mathcal{O}(m^{-3})$, is then

$$\tilde{\mathcal{H}} = mc^2 + V + \frac{\boldsymbol{\pi}^2}{2m} + \frac{e\hbar}{2mc}\,\boldsymbol{B}\cdot\boldsymbol{\sigma} + \frac{\hbar^2}{8m^2c^2}\,\boldsymbol{\nabla}^2 V + \frac{\hbar}{4m^2c^2}\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\,V\times\boldsymbol{\pi}\ . \tag{4.336}$$