



Mathematics and Statistics Doctoral Training Centre

Analysis of the Tight Binding Model

Summer Research Project

Jack Thomas

Supervised by Christoph Ortner



University of Warwick

August 2018

Abstract

We apply the super-cell method to define Fermi levels on finite sub-lattices within the framework of the canonical ensemble and tight binding model. These Fermi levels of course depend on the size of the sub-lattice and the model, but also on temperature via the Fermi-Dirac distribution. We show that the zero temperature and thermodynamic limits of the Fermi levels exist and commute. We then perturb the system by considering more general configurations of the lattice and find that, in the case of metals, the aforementioned results still hold. In the more complex case of insulators, we note that the results break down due to the introduction of new eigenstates within the band gap.

Roscoff, France

July 2018

Acknowledgements

Firstly, I would like to thank my supervisor, Prof. Christoph Ortner, for agreeing to supervise this project. He has provided endless enthusiasm, support and patience throughout. I am very much looking forward to continuing to work together on the PhD project.

I would also like to thank the following people

- Everyone in the office: Hanson, Dom, Paolo, Trish, Phil, Jacek, Jaro, Chris and Osian — it's been a joy struggling through this intense MSc year with you!
- The MASDOC lecturers and staff for the enthusiastic teaching of the modules and for providing a fantastic working environment.
- Students and staff at the ISCD Summer School on Scientific Trends at the Interfaces: Mathematics — Chemistry — High Performance Computing. Special thanks go to Eric Cancès, David Gontier and Geneviève Dusson for their time and expertise.

This work was supported by EPSRC as part of the Mathematics and Statistics Doctoral Training Centre at the University of Warwick, Grant No. EP/H023364/1.

Contents

Abstract	i
Acknowledgements	ii
Declarations	iv
1 Tight Binding Approximation	1
2 Homogeneous Lattice	4
2.1 Definitions & Results	4
2.2 Examples: Metals	9
2.3 Examples: Insulators	12
3 Defects	15
3.1 Definitions & Notation	15
3.2 Results	17
4 Proofs	22
4.1 Proof of Theorem 1	23
4.2 Proof of Theorem 2	33
Bibliography	v

Declarations

To the best of my knowledge, I have presented only my original work and the outcomes of my collaboration with Prof. Christoph Ortner, unless specifically referenced or cited in the bibliography. This work has not been submitted for another degree or qualification at the University of Warwick or any other higher-education institution.

Chapter 1

Tight Binding Approximation

In this section we give a very brief introduction to the formulation of the tight binding model, largely following [9]. Starting with the time independent Schrödinger equation, we will explain the assumptions made in order to arrive at the tight binding model. For a more detailed exposition, see [11, 13, 16, 20].

We shall consider a system consisting of M nuclei positioned at $R_1, \dots, R_M \in \mathbb{R}^3$ and N electrons at $r_1, \dots, r_N \in \mathbb{R}^3$. We shall make the *Born-Oppenheimer approximation* and assume that the nuclei positions are fixed. This is a good approximation in most circumstances [9] and is based upon the fact that nuclei move very slowly relative to the electrons in most cases.

The Hamiltonian is an operator, $\hat{\mathcal{H}}$, describing the total energy of the system. This is composed of the kinetic energy and the potential energy operators. Since, we are considering the nuclei as fixed, the kinetic energy operator is of the form

$$\sum_{j=1}^N -\frac{\hbar^2}{2m} \Delta_j$$

where $\hbar$ is the reduced Planck constant and m is the mass of the electron. On the other hand, the energy due to the interactions between the electrons and nuclei is expressed in the potential energy term:

$$\frac{1}{2} \sum_{\substack{1 \leq j, k \leq N \\ j \neq k}} \frac{e^2}{4\pi\epsilon_0 |r_j - r_k|} + \sum_{\substack{1 \leq j \leq M \\ 1 \leq k \leq N}} \frac{-e^2 Z_j}{4\pi\epsilon_0 |r_k - R_j|}$$

where e is the electronic charge, Z_j is the atomic number of atom the atom positioned at R_j and ϵ_0 is the permittivity of free space. This corresponds to the electron-electron and

the electron-nuclei interactions respectively. The time-independent Schrödinger equation takes the form of the eigenvalue problem for the operator $\hat{\mathcal{H}}$:

$$\hat{\mathcal{H}}\Psi = E\Psi. \quad (1.1)$$

We call the Ψ satisfying this equation the many-body wave functions and E the corresponding energies. Notice that, for fixed nuclei positions, the complex-valued wave functions only depend upon the position and spin of the electrons. The wave functions are normalised so that they give rise to probability densities $|\Psi|^2$ which represents the probability of finding the system in a given state [20].

Further, we shall make the *mean field approximation* in which the electron-electron interactions are replaced by a term representing the average potential generated by the electrons [9]. This is then added to the electron-nuclei potential giving an *effective potential* which we will denote by $\hat{V}_{\text{eff}}$. Now, the Hamiltonian can be written as the sum of single-particle contributions, $\hat{\mathcal{H}}_j$, and thus (1.1) becomes

$$\sum_{j=1}^N \hat{\mathcal{H}}_j \Psi = \sum_{j=1}^N \left(-\frac{\hbar^2}{2m} \Delta_j + V_{\text{eff}}(r_j) \right) \Psi = E\Psi.$$

For the following, we shall denote by $\hat{\mathcal{H}}^{sp}$ the single-particle Hamiltonian. Suppose we are given a set of *basis* functions $\{\phi_\alpha(\cdot - R_j)\}_\alpha$ centred on each atom j . We assume that we can approximate the solutions to the single-particle Schrödinger equation by linear combination of our basis functions. That is, if ψ is some solution to the single-particle Schrödinger equation with corresponding energy, ε , then there is some constants $C_{j\alpha}$ such that

$$\psi(r) = \sum_{j\alpha} C_{j\alpha} \phi_\alpha(r - R_j).$$

Now, since ψ is a solution to the single-particle equation, we know that

$$\sum_{j\alpha} C_{j\alpha} \hat{\mathcal{H}}^{sp} \phi_\alpha(r - R_j) = \varepsilon \sum_{j\alpha} C_{j\alpha} \phi_\alpha(r - R_j).$$

After multiplying on the left by some $\phi_\beta(\cdot - R_k)$ and integrating over space, we obtain the following matrix equation:

$$\mathcal{H}C = \varepsilon SC$$

where $\mathcal{H}$ and S are the Hamiltonian and overlap matrices respectively:

$$\mathcal{H}_{j\alpha, k\beta} := \int \phi_\beta(\cdot - R_k) \hat{\mathcal{H}}^{sp} \phi_\alpha(\cdot - R_j) \quad \text{and} \quad S_{j\alpha, k\beta} := \int \phi_\beta(\cdot - R_k) \phi_\alpha(\cdot - R_j).$$

This matrix equation is known as the *Kohn-Sham* equation and of course depends upon the choice of basis [9]. In the tight binding framework, a minimal set of basis functions consisting of atomic-like orbitals is chosen. These functions are localised (*tightly bound*) around the atomic positions and so off-diagonal entries of the overlap matrix are small. In this report, we will only consider *orthogonal tight binding* where the overlap matrix is assumed to be the identity. This assumption can be justified by replacing the atomic orbitals by Löwdin functions [19, 15].

After making the simplifying assumption that the effective potential can be written as a sum of contributions centred on each atom:

$$V_{\text{eff}}(r) = \sum_j V_{\text{eff}}^{(j)}(r - R_j),$$

we can see that the Hamiltonian matrix elements are composed of on-site terms and two- and three-centre integrals. On-site terms only depend upon a single atomic position and are of the form

$$\int \phi_\beta(\cdot - R_j) \left(-\frac{1}{2}\Delta + V_{\text{eff}}^{(j)}(\cdot - R_j) \right) \phi_\alpha(\cdot - R_j).$$

On the other hand, taking $j \neq k$, we have two-centre contributions of the form

$$\begin{aligned} & \int \phi_\beta(\cdot - R_k) \left(-\frac{1}{2}\Delta \right) \phi_\alpha(\cdot - R_j) \quad \text{and} \\ & \int \phi_\beta(\cdot - R_k) V_{\text{eff}}^{(j)}(\cdot - R_j) \phi_\alpha(\cdot - R_j) \end{aligned}$$

which only depend on the vector between R_k and R_j and three-centre contributions of the form

$$\sum_{l \neq j, k} \int \phi_\beta(\cdot - R_k) \left(V_{\text{eff}}^{(l)}(\cdot - R_l) \right) \phi_\alpha(\cdot - R_j).$$

Since we have these integrals between each pair of atoms and each combination of atomic orbitals, the number of integrals is very large for large systems. Even if we neglect three-centre integrals, which is done in a number of tight binding models [9], and consider N_b orbitals, the number of Hamiltonian entries between each pair of atoms is N_b^2 . It can be shown however [19], that the number of variables can be drastically reduced by considering the various symmetries of the atom-like orbitals.

Chapter 2

Homogeneous Lattice

2.1 Definitions & Results

We shall consider the infinite homogeneous lattice $\Lambda := \mathbb{Z}^d$ and finite subsystems $\Lambda_R := \{-R + 1, \dots, R\}^d \subset \Lambda$ with periodic boundary conditions. This is for simplicity of notation and the same results hold in the more general case of Bravais lattices without any complication.

We will denote by Ξ a finite set of n_Ξ atomic orbitals and thus, the Hamiltonian matrix for the finite subsystem Λ_R will be a $\Xi \times \Xi$ block matrix consisting of $(2R)^d \times (2R)^d$ blocks. We further interpret $(\psi^\alpha(\ell))_{\ell \in \Lambda_R}^{\alpha \in \Xi}$ as a vector in the following way:

$$\mathcal{H}^R = \begin{pmatrix} \mathcal{H}^{\alpha_1 \alpha_1} & \dots & \mathcal{H}^{\alpha_1 \alpha_{n_\Xi}} \\ \vdots & \ddots & \vdots \\ \mathcal{H}^{\alpha_{n_\Xi} \alpha_1} & \dots & \mathcal{H}^{\alpha_{n_\Xi} \alpha_{n_\Xi}} \end{pmatrix} \quad \text{and} \quad (\psi^\alpha(\ell))_{\ell \in \Lambda_R}^{\alpha \in \Xi} = \begin{pmatrix} (\psi^{\alpha_1}(\ell))_\ell \\ \vdots \\ (\psi^{\alpha_{n_\Xi}}(\ell))_\ell \end{pmatrix}.$$

Thus, we have the following matrix-vector multiplication:

$$(\mathcal{H}^R \psi)_\ell^\alpha = \sum_{\beta \in \Xi} \sum_{k \in \Lambda_R} \mathcal{H}_{\ell k}^{\alpha \beta} \psi^\beta(k).$$

We make the following standard assumptions on the Hamiltonian:

Assumption (Sym). *The Hamiltonian is symmetric: for each $\ell, k \in \Lambda_R$ and $\alpha, \beta \in \Xi$*

$$\mathcal{H}_{\ell k}^{\alpha \beta} = \mathcal{H}_{k \ell}^{\beta \alpha}.$$

Assuming the following *two-centre approximation* will allow us to explicitly calculate eigenvalues of the Hamiltonian:

Assumption (2C). *There exist functions $h^{\alpha\beta} \in \ell^1(\Lambda)$ such that, for each $\alpha, \beta \in \Xi$ and $\ell, k \in \Lambda$ we have*

$$\mathcal{H}_{\ell k}^{\alpha\beta} = h^{\alpha\beta}(k - \ell).$$

The *continuous* and *discrete Brillouin zones* associated with, respectively, Λ and Λ_R , will be important for discussing the energy bands (defined in Proposition 1) of the system. These are, respectively, defined by

$$\text{BZ} := (-\pi, \pi]^d \quad \text{and} \quad \text{BZ}_R := \left\{ -\pi + \frac{j\pi}{R} : j = 1, \dots, 2R \right\}^d \subset \text{BZ}.$$

Under the above assumptions, we can obtain explicit formulae for the energy bands of the system:

Proposition 1. *Suppose that our Hamiltonian matrices on Λ satisfies (TB), (Sym) and (2C). Then, there exists n_{Ξ} continuous functions $\varepsilon^{\alpha} : \overline{\text{BZ}} \rightarrow \mathbb{R}$ for $\alpha \in \Xi$ such that*

$$\sigma(\mathcal{H}^R) = \{\varepsilon^{\alpha}(k) : k \in \text{BZ}_R, \alpha \in \Xi\}.$$

We shall call the functions $\varepsilon^{\alpha}(\cdot)$ energy bands corresponding to the Hamiltonian. Moreover, the energy bands are given as the n_{Ξ} roots of the equation

$$\det \left(\hat{h}(k) - \varepsilon(k) \text{Id}_{\Xi} \right) = 0$$

where $\hat{h}(k)$ is the $n_{\Xi} \times n_{\Xi}$ matrix composed of the Fourier transforms $\widehat{h^{\alpha\beta}}(k)$, for each $\alpha, \beta \in \Xi$.

Remark 1. *In general, the labelling of the energy bands is not unique. The non-uniqueness results from the fact that energy bands may intersect. We will fix some choice of continuous energy bands $\varepsilon^{\alpha} : \text{BZ} \rightarrow \mathbb{R}$ for $\alpha \in \Xi$. These maps are of course real-valued by the symmetry of the Hamiltonian.*

By Proposition 1, we know that the spectrum, $\sigma(\mathcal{H}^R)$, is uniformly bounded in $\mathbb{R}$. Thus, for the remainder of this report, we shall fix $[\underline{\sigma}, \overline{\sigma}]$ such that $\sigma(\mathcal{H}^R) \subset [\underline{\sigma}, \overline{\sigma}]$.

We now introduce the Fermi-Dirac distribution $f_{\beta}(\cdot)$ given by

$$f_{\beta}(E) = \frac{1}{1 + e^{\beta E}}$$

which gives the probability of an electron occupying a particular energy level [9]. The parameter $\beta > 0$ is equal to $(k_B T)^{-1}$ where k_B is the Boltzmann constant and T is the absolute temperature. We will consider the zero temperature limit ($\beta \rightarrow \infty$) and so we note that

$$f_\infty(E) := \lim_{\beta \rightarrow \infty} f_\beta(E) = \begin{cases} 1 & \text{if } E < 0 \\ \frac{1}{2} & \text{if } E = 0 \\ 0 & \text{if } E > 0. \end{cases}$$

For a given *chemical potential* τ , we can define the local particle number for each finite subsystem ($R < \infty$) by

$$N_{\#}^{\beta, R}(\tau) := \frac{2}{\#\sigma(\mathcal{H}^R)} \sum_{\alpha \in \Xi} \sum_{k \in \text{BZ}_R} f_\beta(\varepsilon^\alpha(k) - \tau)$$

for each $\beta \in (0, \infty]$. As shown in [5] for the local particle number and other local observables, we can consider the thermodynamic limit ($R \rightarrow \infty$) of the local particle number to obtain the local particle number for the infinite crystal, denoted by $N_{\#}^{\beta, \#}(\cdot)$. In the case of the two-centre approximation that we are considering, we obtain

$$N_{\#}^{\beta, \#}(\tau) = \frac{2}{n_{\Xi}} \sum_{\alpha \in \Xi} \frac{1}{|\text{BZ}|} \int_{\text{BZ}} f_\beta(\varepsilon^\alpha(k) - \tau) dk.$$

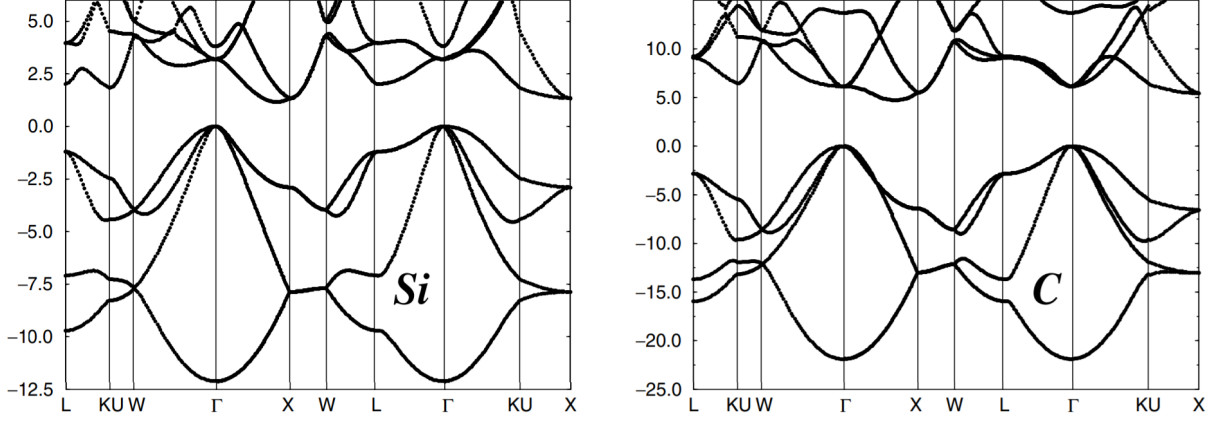
It will be useful to consider ρ_R and the *density of states*, ρ , which are probability measures defined by

$$\begin{aligned} \rho_R(A) &:= \frac{1}{n_{\Xi}} \sum_{\alpha \in \Xi} \frac{1}{\#\text{BZ}_R} \#\{k \in \text{BZ}_R : \varepsilon^\alpha(k) \in A\} \quad \text{and} \\ \rho(A) &:= \frac{1}{n_{\Xi}} \sum_{\alpha \in \Xi} \frac{1}{|\text{BZ}|} |\{k \in \text{BZ} : \varepsilon^\alpha(k) \in A\}| \end{aligned}$$

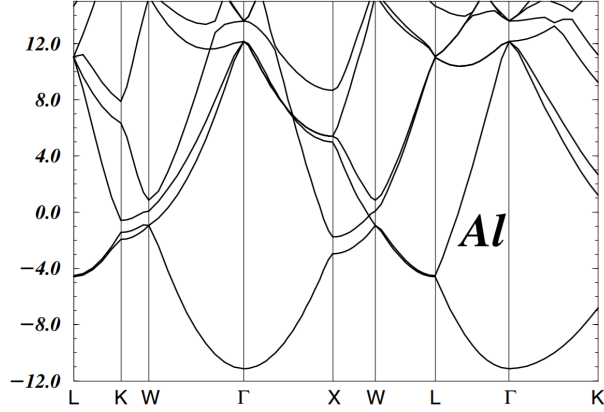
for every Borel $A \subset \mathbb{R}$. These definitions lead to simple formulae for the local particle number functions:

$$N_{\#}^{\beta, R}(\tau) = 2 \int f_\beta(\cdot - \tau) d\rho_R \quad \text{and} \quad N_{\#}^{\beta, \#}(\tau) = 2 \int f_\beta(\cdot - \tau) d\rho.$$

Remark 2. *It is important to note that the spectrum corresponding to each finite subsystem is contained within the support of the density of states. Indeed, if we take some eigenvalue of $\mathcal{H}^R$ then, as shown in Proposition 1, it is of the form $\varepsilon^\alpha(k)$ for some $\alpha \in \Xi$ and $k \in \text{BZ}_R$. By the continuity of $\varepsilon^\alpha(\cdot)$ on BZ , we necessarily have $\rho(N) > 0$ for all open sets N that contain $\varepsilon^\alpha(k)$.*



(a) Band structure of Silicone (left), a semiconductor, and Carbon (right), an insulator.



(b) Band structure of Aluminium, a metal.

Figure 2.1: Band structure of typical semiconductors, insulators and metals [13]. Here, the symbols on the x-axis represent high symmetry points in the Broullin zone.

We now make the following assumption on the energy bands:

Assumption (E). *There exists a constant $M < \infty$ such that, for all $\alpha \in \Xi$ and $\tau \in \mathbb{R}$, the level sets $\{k \in \text{BZ} : \varepsilon^\alpha(k) = \tau\}$ contain at most M points.*

In particular, (E) asserts that the energy bands are not constant on any open ball contained in BZ. Therefore, for all $\tau \in \mathbb{R}$, we have that $\rho(\{\tau\}) = 0$. As we shall see, this assumption is satisfied for all the mathematical examples we consider in this report. Moreover, the energy bands resulting from real materials satisfies this assumption in a large number of cases. This is demonstrated in Figure 2.1 for typical insulators, metals and semiconductors.

We shall consider the canonical ensemble in which the number of electrons are fixed. In this case, the Fermi level is defined so that we have the correct number of electrons in the system [9]:

Definition 1 (Fermi Level for finite temperature). *Fix some $0 < \beta < \infty$ and $R \in \mathbb{N}$. We define $\varepsilon_F^{\beta,R}$ and $\varepsilon_F^{\beta,\#}$ to be the unique solutions of the non-linear equations*

$$N_{\#}^{\beta,R}(\varepsilon_F^{\beta,R}) = 1 \quad \text{and} \quad N_{\#}^{\beta,\#}(\varepsilon_F^{\beta,\#}) = 1 \quad \text{respectively.}$$

We call $\varepsilon_F^{\beta,R}$ (and $\varepsilon_F^{\beta,\#}$) the Fermi Level for $\beta < \infty$ and the finite system Λ_R (and for the infinite lattice Λ respectively).

Remark 3. *Note that, $N_{\#}^{\beta,R}$ and $N_{\#}^{\beta,\#}$ are strictly increasing continuous functions and thus have strictly increasing continuous inverses. Furthermore, $N_{\#}^{\beta,R}(\tau) \rightarrow 0$ as $\tau \rightarrow -\infty$ and $N_{\#}^{\beta,R}(\tau) \rightarrow 2$ as $\tau \rightarrow \infty$ and so Definition 1 is valid.*

Now, if it is possible, we similarly define the zero temperature Fermi Level for the infinite crystal:

Definition 2 (Fermi level for zero temperature). *In the case that there exists a unique $\tau \in \mathbb{R}$ satisfying*

$$N_{\#}^{\infty,\#}(\tau) = 1,$$

we denote this τ by $\varepsilon_F^{\infty,\#}$ and call it the zero temperature Fermi Level for the infinite lattice Λ .

Definition 2 asserts that, at zero temperature, exactly half of the total number of electronic states are filled [4].

Notice that, if the Definition 2 is not satisfied for a unique value, then there exists an interval $[\underline{\varepsilon}, \bar{\varepsilon}]$ of positive length such that $N_{\#}^{\infty,\#} \equiv 1$ on $[\underline{\varepsilon}, \bar{\varepsilon}]$. The maximal such interval is known as the *band gap*.

The first main aim of this report is to prove the following theorem:

Theorem 1 (Fermi levels commute). *Suppose that, for each $R \in \mathbb{N}$, we have a Hamiltonian $\mathcal{H}^R$ satisfying **(Sym)** and **(2C)**. Moreover, suppose that the corresponding energy bands satisfy **(E)**. Then, the zero temperature and thermodynamic limits of the Fermi levels commute and satisfy*

$$\lim_{R \rightarrow \infty} \lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta,R} = \lim_{\beta \rightarrow \infty} \lim_{R \rightarrow \infty} \varepsilon_F^{\beta,R} = \varepsilon_F^{\infty,\#} := \frac{1}{2}(\underline{\varepsilon} + \bar{\varepsilon}).$$

Furthermore, we may take the diagonal limit,

$$\lim_{R \rightarrow \infty} \varepsilon_F^{\beta_R, R} = \varepsilon_F^{\infty, \#},$$

for any sequence $(\beta_R)_R$ such that $\beta_R > 0$ for all R and $\beta_R \rightarrow \infty$ as $R \rightarrow \infty$.

Theorem 1 allows us to extend the definition of $\varepsilon_F^{\infty, \#}$ to the case where we have a band gap (i.e. when Definition 2 no longer applies). In Sections 2.2 and 2.3, we shall discuss examples of this result and a full proof will be given in Chapter 4.

2.2 Examples: Metals

We first consider the case of metals where there is no band gap in the thermodynamic limit of the spectrum. This means that the Fermi level on the infinite crystal at zero temperature is well defined. Throughout this section, we will assume that there is only one atomic orbital per atom. This is for simplicity of notation but does not affect the results described.

We will now see the calculation of $\varepsilon_F^{\beta, R}$ and $\varepsilon_F^{\beta, \#}$ for a simple two centre approximation model that only accounts for nearest neighbour interactions:

Example 1. Denoting by e_j the standard basis vectors of $\mathbb{R}^d$, we assume that the Hamiltonian matrix takes the form:

$$h(\ell) = \begin{cases} h_0 & \text{if } \ell = 0 \\ h_1 & \text{if } \ell = \pm e_j \text{ for some } j = 1, \dots, d \\ 0 & \text{otherwise.} \end{cases}$$

As shown in Proposition 1, the eigenvalues are given by:

$$\varepsilon(k) = \hat{h}(k) = \sum_{\ell \in \Lambda_R} h(\ell) e^{i\ell \cdot k} = h_0 + 2h_1 \sum_{j=1}^d \cos(k_j). \quad (2.1)$$

for $k \in \text{BZ}_R$. We will show that this means $\varepsilon_F^{\beta, R} = h_0$ for all finite $\beta > 0$ and $R > 0$.

Proof. Without loss of generality, we may assume that $h_0 = 0$. If not, we can always shift the Hamiltonian so that the diagonal entries are zero. We note that by (2.1), it is

clear that $\sigma(\mathcal{H}^R)$ is symmetric about 0. Moreover, since $\#\sigma(\mathcal{H}^R) = (2R)^d$ is even, we can conclude that we have an even number of zero eigenvalues. This, along with the fact that $f_\beta(E) + f_\beta(-E) = 1$ for all E and $\beta \in (0, \infty]$, means that we have

$$\sum_{\varepsilon \in \sigma(\mathcal{H}^R)} f_\beta(\varepsilon) = \frac{(2R)^d}{2}$$

and so $\varepsilon_F^{\beta,R} = 0$ for each β and finite R .

Now, for the thermodynamic limit, we will show that $\varepsilon_F^{\beta,\#} = 0$. By the periodicity of the cosine function, we have

$$\int_{(-\pi,\pi]^d} f_\beta(2h_1 \sum_j \cos(k_j)) dk = 2^d \int_{(0,\pi]^d} f_\beta(2h_1 \sum_j \cos(k_j)) dk$$

After defining $g(k) = f_\beta(2h_1 \sum_j \cos(k_j))$, we have $g(k) = 1 - g(\pi \mathbf{1} - k)$ for each $k \in (0, \pi]^d$ where $\mathbf{1} = (1, \dots, 1)$. This means that the area inside the box $(0, \pi]^d \times [0, 1]$ bounded above by the surface $y_{d+1} = g(k)$ is equal to that bounded below by the same surface. We therefore obtain

$$\int_{(0,\pi]^d} f_\beta(2h_1 \sum_j \cos(k_j)) dk = \pi^d - \int_{(0,\pi]^d} f_\beta(2h_1 \sum_j \cos(k_j)) dk = \frac{\pi^d}{2}.$$

This means that the original integral is equal to $\frac{(2\pi)^d}{2}$ and so $\varepsilon_F^{\beta,\#} = 0$. We have thus showed that $\varepsilon_F^{\beta,R} = \varepsilon_F^{\beta,\#} = 0$ for all $R > 0, \beta \in (0, \infty]$. $\square$

For this particular example, firstly we have shown that $\varepsilon_F^{\beta,R} = h_0$ for all β, R and thus, trivially, $\varepsilon_F^{\beta,R}$ converges in the thermodynamic limit. Moreover, we have seen that the limiting value is $\varepsilon_F^{\beta,\#}$. Also, in the case of Example 1, it is trivial to see that the limits $\lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta,\#}$ and $\lim_{R \rightarrow \infty} \varepsilon_F^{\beta_R,R}$ (for any choice of $\beta_R \rightarrow \infty$) exist. This is in agreement with the claims made in Theorem 1.

In addition, Theorem 1 asserts that $\lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta,\#} = \varepsilon_F^{\infty,\#}$ which, since we are assuming there is only one energy band, is given by Definition 2. For the case of Example 1, we have $\varepsilon_F^{\beta,\#} = h_0$ for all β , and so we also have $\varepsilon_F^{\infty,\#} = h_0$.

To complete this section, we look at the case where the Hamiltonian is given by a two-centre approximation and show an example where the zero temperature Fermi Level for the infinite lattice need not be h_0 . This also means that the zero temperature limit of $\varepsilon_F^{\beta,\#}$ need not be h_0 for a general two-centre approximation model.

Example 2. In this example, we consider $d = 1$ and the Hamiltonian given by

$$h(\ell) = \begin{cases} h_0 & \text{for } \ell = 0 \\ h_1 & \text{for } \ell = \pm 1 \\ h_2 & \text{for } \ell = \pm 2 \\ 0 & \text{otherwise.} \end{cases}$$

In the following we will take $h_1 = h_2 = \frac{1}{2}$ for simplicity and, without loss of generality, take $h_0 = 0$. In this case, the spectrum in the thermodynamic limit is given by

$$\varepsilon(k) = \cos(k) + \cos(2k) \quad \text{for } k \in \text{BZ}. \quad (2.2)$$

Because we only have one energy band, $\varepsilon_{\text{F}}^{\infty, \#}$ is well defined by Definition 2. Recall, this is the τ satisfying

$$\frac{\pi}{2} = \int_0^\pi f_\infty(\varepsilon(k) - \tau) dk = |\{k \in [0, \pi] : \cos(k) + \cos(2k) < \tau\}|.$$

Let us find points $k \in [0, \pi]$ such that $\varepsilon(k) = \tau$. By applying the double angle formula and (2.2), we know that $\varepsilon(k) = \tau$ is equivalent to

$$2 \cos^2(k) + \cos(k) - (\tau + 1) = 0.$$

If this equation had no roots then, by continuity, $\cos(k) + \cos(2k) - \tau$ would be always positive or always negative. This clearly means τ is not the required Fermi level. Solving this quadratic equation in $\cos(k)$, we have

$$\cos(k) = \frac{-1}{4} \pm \frac{1}{4} \sqrt{D_\tau} \quad \text{where } D_\tau = 8\tau + 9.$$

Supposing that $-\frac{1}{4} \pm \frac{1}{4} \sqrt{D_\tau} \in [-1, 1]$, allows us to find $k_\tau^\pm \in [0, \pi]$ such that

$$\cos(k_\tau^\pm) = -\frac{1}{4} \pm \frac{1}{4} \sqrt{D_\tau}$$

Now, $\cos(k) + \cos(2k) < \tau$ exactly on the interval (k_τ^+, k_τ^-) and so

$$|\{k \in [0, \pi] : \varepsilon(k) < \tau\}| = k_\tau^- - k_\tau^+.$$

In order to find the Fermi level, we therefore wish to solve $k_\tau^- = k_\tau^+ + \frac{\pi}{2}$. Taking cosines of both sides and using the fact that $\cos(x + \pi/2) = -\sin(x)$, we have

$$\begin{aligned} -\frac{1}{4} - \frac{1}{4} \sqrt{D_\tau} &= -\sin\left(\cos^{-1}\left(\frac{-1}{4} + \frac{1}{4} \sqrt{D_\tau}\right)\right) \\ &= -\sqrt{1 - \left(\frac{-1}{4} + \frac{1}{4} \sqrt{D_\tau}\right)^2}. \end{aligned}$$

Rearranging this yields $D_\tau = 7$ and $\tau = -\frac{1}{4}$. Since $\frac{-1}{4} \pm \frac{1}{4}\sqrt{7} \in [-1, 1]$, the calculations above are valid and the Fermi level that we are interested in is $\varepsilon_F^{\infty, \#} = -\frac{1}{4}$.

2.3 Examples: Insulators

As shown in Section 2.2, for a band gap to be present, we will need to consider more than one atomic orbital per atom. We will now see an explicit example to show that the energy bands need not overlap and so band gaps are possible.

Example 3 (Two orbitals per atom). *This example is adapted from [13] and [20]. Firstly, we shall consider the finite 1 dimensional system $\Lambda_R := \{-R+1, \dots, R\}$ endowed with periodic boundary conditions. Each atom in the system will be associated with two atomic orbitals. We shall consider the 5-parameter Hamiltonian matrix given by a simple nearest neighbour model. That is, if we use the notation $\ell \sim_{\Lambda_R} k$ to mean ℓ and k are nearest neighbours in Λ_R then, for all $\ell, k \in \Lambda_R$ and $\alpha, \beta \in \{1, 2\}$, we have*

$$\mathcal{H}_{\ell k}^{\alpha\beta} = \begin{cases} h_0^\alpha & \text{if } \ell = k \text{ and } \alpha = \beta \\ h_1^\alpha & \text{if } \ell \sim_{\Lambda_R} k \text{ and } \alpha = \beta \\ h_1^{12} & \text{if } \ell \sim_{\Lambda_R} k \text{ and } \alpha \neq \beta \\ 0 & \text{otherwise} \end{cases}$$

for some constants $h_0^1, h_0^2, h_1^1, h_1^{12}, h_1^2$. Because of the translational symmetry in the system, we may apply the calculation in Proposition 1 to conclude, for each $k \in \text{BZ}_R$, the eigenvalues are solutions to the equation

$$|\hat{h}(k) - \varepsilon(k)I| = 0.$$

A simple calculation reveals

$$\hat{h}(k) = \begin{pmatrix} h_0^1 + 2h_1^1 \cos(k) & 2h_1^{12} \cos(k) \\ 2h_1^{12} \cos(k) & h_0^2 + 2h_1^2 \cos(k) \end{pmatrix}$$

and thus we obtain the quadratic in $\varepsilon(k)$:

$$\begin{aligned} \varepsilon(k)^2 - \varepsilon(k) (h_0^1 + 2h_1^1 \cos(k) + h_0^2 + 2h_1^2 \cos(k)) + \\ + (h_0^1 + 2h_1^1 \cos(k)) (h_0^2 + 2h_1^2 \cos(k)) - (2h_1^{12} \cos(k))^2 = 0. \end{aligned}$$

After denoting $a(k) := h_0^1 + 2h_1^1 \cos(k)$ and $b(k) := h_0^2 + 2h_1^2 \cos(k)$, we see that the discriminant of this quadratic is

$$(a(k) + b(k))^2 - 4a(k)b(k) + 16(h_1^{12} \cos(k))^2 = (a(k) - b(k))^2 + 16(h_1^{12} \cos(k))^2$$

and so we see that there are two energy bands given by:

$$\varepsilon^\pm(k) = \frac{a(k) + b(k)}{2} \pm \sqrt{\frac{(a(k) - b(k))^2}{4} + 4(h_1^{12} \cos(k))^2}.$$

It is easy to see that, for $k \in \text{BZ}_R$, these bands do not necessarily intersect as is the case of Figure 2.2.

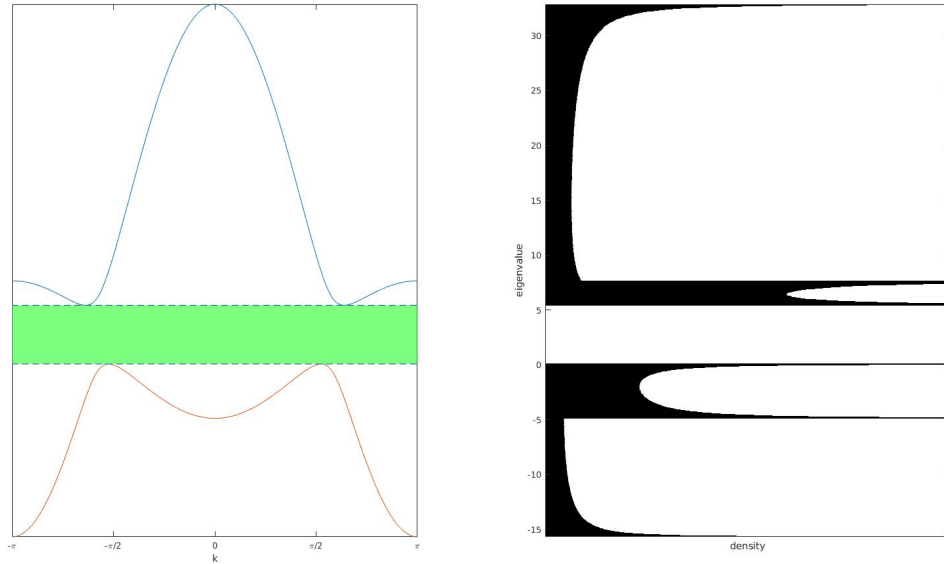


Figure 2.2: Energy bands resulting from the simple 5-parameter model from Example 3 with constants $h_0^1 = 0, h_0^2 = 10, h_1^1 = -1, h_1^2 = 10$ and $h_1^{12} = 5$. Left: the two energy bands with the band gap shown in green. Right: associated approximate density of states.

For the remainder of this report, we will denote the band gap by $[\underline{\varepsilon}, \bar{\varepsilon}]$. Since $N_{\#}^{\infty, \#} \equiv 1$ on the whole band gap, we cannot use $N_{\#}^{\infty, \#}(\cdot)$ alone to define $\varepsilon_{\text{F}}^{\infty, \#}$. We will instead show that the zero temperature limit of $\varepsilon_{\text{F}}^{\beta, \#}$ exists and thus define $\varepsilon_{\text{F}}^{\infty, \#} := \lim_{\beta \rightarrow \infty} \varepsilon_{\text{F}}^{\beta, \#}$. Moreover, we show that this limiting value is the midpoint of the band gap and is therefore consistent with the definition of $\varepsilon_{\text{F}}^{\infty, \#}$ given in the case of a metal (Definition 2).

The results described in this chapter can be summarised by the commutative diagram in Figure 2.3.

$$\begin{array}{ccc}
 \varepsilon_F^{\beta,R} & \xrightarrow{\quad\quad\quad} & \varepsilon_F^{\beta,\#} \\
 \downarrow & \searrow \text{~~~~~} & \downarrow \\
 \lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta,R} & \longrightarrow & \varepsilon_F^{\infty,\#} := \frac{1}{2} (\underline{\varepsilon} + \bar{\varepsilon})
 \end{array}$$

Figure 2.3: Commutative diagram representing Theorem 1. The horizontal arrows represent the thermodynamic limits ($R \rightarrow \infty$) and the zero temperature limits ($\beta \rightarrow \infty$) are represented by the vertical arrows. The diagonal arrow represents the limit $\lim_{R \rightarrow \infty} \varepsilon_F^{\beta_R,R}$ for any choice of sequence of positive numbers $\beta_R \rightarrow \infty$ as $R \rightarrow \infty$.

Chapter 3

Defects

3.1 Definitions & Notation

We shall now consider inhomogeneous lattice systems. This is equivalent to perturbing the atoms of the homogeneous lattice by some displacement $u: \Lambda \rightarrow \mathbb{R}^d$. We shall see how this perturbation affects the results described in Chapter 2.

Let $\mathcal{V}_{\mathfrak{m}}$ denote the set of all *admissible configurations* $y: \Lambda \rightarrow \mathbb{R}^d$ satisfying the following non-interpenetration condition:

$$|y(\ell) - y(k)| \geq \mathfrak{m}|\ell - k| \quad \text{for all } \ell, k \in \Lambda$$

for the positive constant $\mathfrak{m} > 0$.

In this setting, the Hamiltonian matrices also depend upon the configuration $y \in \mathcal{V}_{\mathfrak{m}}$ and the assumptions that we impose now take the following form:

Assumption (Sym). *We further suppose that the Hamiltonian is symmetric: for all $y \in \mathcal{V}_{\mathfrak{m}}$*

$$\mathcal{H}_{\ell k}^{\alpha\beta}(y) = \mathcal{H}_{k\ell}^{\beta\alpha}(y)$$

for each $\ell, k \in \Lambda_R$ and atomic orbitals $\alpha, \beta \in \Xi$.

Assumption (2C). *There exist functions $h^{\alpha\beta}$ such that*

$$\mathcal{H}_{\ell k}^{\alpha\beta}(y) = h^{\alpha\beta}(y(k) - y(\ell))$$

for each $\alpha, \beta \in \Xi$ and $\ell, k \in \Lambda$.

We introduce the following additional assumption on the Hamiltonian:

Assumption (TB). *The functions $h^{\alpha\beta}$ are continuously differentiable and there exist positive constants $h_0, \gamma_0, h_1, \gamma_1 > 0$ such that, for all $y \in \mathcal{V}_m$*

$$\begin{aligned} |h^{\alpha\beta}(y(\ell) - y(k))| &\leq h_0 e^{-\gamma_0 |y(\ell) - y(k)|} \quad \text{and} \\ |\nabla h^{\alpha\beta}(y(\ell) - y(k))| &\leq h_1 e^{-\gamma_1 |y(\ell) - y(k)|} \end{aligned}$$

for all $\ell, k \in \Lambda$ and $\alpha, \beta \in \Xi$.

Under these assumptions we again obtain the uniform boundedness of the spectrum:

Proposition 2 (Lemma 2.1 [6]). *Under the assumptions (2C), (TB) and (Sym), the spectrum of $\mathcal{H}^R(y)$ is uniformly bounded in $\mathbb{R}$ independent of R and the configuration $y \in \mathcal{V}_m$. That is, there exists $\underline{\sigma}, \bar{\sigma}$ such that $\sigma(\mathcal{H}^R(y)) \subset [\underline{\sigma}, \bar{\sigma}]$ for all R and $y \in \mathcal{V}_m$.*

It will be convenient to consider *displacements* $u: \Lambda \rightarrow \mathbb{R}^d$ such that the configuration $x + u \in \mathcal{V}_m$ where $x: \Lambda \rightarrow \Lambda$ is the identity. For each $\gamma > 0$, we define the semi-norm, $\|\cdot\|_\gamma$, on the space of displacements by

$$\|u\|_\gamma^2 := \sum_{\ell, k \in \Lambda} e^{-2\gamma|\ell-k|} |u(\ell) - u(k)|^2.$$

All of these semi-norms are equivalent, [7], and so we can define the set of *admissible displacements* by

$$\text{Adm}_m := \{u: \Lambda \rightarrow \mathbb{R}^d: \|u\|_\gamma < \infty, \text{ and } x + u \in \mathcal{V}_m\}.$$

For a given displacement $u \in \text{Adm}_m$ and chemical potential $\tau \in \mathbb{R}$, we can define the local particle number for each finite subsystem

$$N_{\#}^{\beta, R}(u; \tau) := \frac{2}{\#\sigma(\mathcal{H}^R(x + u))} \sum_{\varepsilon \in \mathcal{H}^R(x + u)} f_{\beta}(\varepsilon - \tau).$$

Thus, we can define the Fermi level as we did for the homogeneous case:

Definition 3. *Fix $\beta \in (0, \infty)$, $R \in \mathbb{N}$ and displacement $u \in \text{Adm}_m$. We define $\varepsilon_F^{\beta, R}(u)$ to be the unique solution of the non-linear equation*

$$N_{\#}^{\beta, R}(u; \varepsilon_F^{\beta, R}(u)) = 1.$$

We call $\varepsilon_F^{\beta, R}(u)$ the *Fermi Level* for $\beta < \infty$ and the finite system Λ_R corresponding to the displacement u .

3.2 Results

Theorem 2. Suppose that $\mathcal{H}^R$ is a Hamiltonian satisfying **(Sym)**, **(2C)** and **(TB)**. Moreover, suppose that the energy bands corresponding to $\mathcal{H}^R(x)$ satisfy **(E)**. For an admissible displacement $u \in \text{Adm}_{\mathbf{m}}$ (for some $\mathbf{m} > 0$) we have

- (1) The thermodynamic limit of the Fermi level coincides with the non-defect case:

$$\lim_{R \rightarrow \infty} \varepsilon_{\mathbf{F}}^{\beta, R}(u) = \varepsilon_{\mathbf{F}}^{\beta, \#} \quad \text{for all } \beta < \infty$$

and thus the zero temperature limit of this quantity is $\varepsilon_{\mathbf{F}}^{\infty, \#}$.

- (2) The zero temperature limit $\lim_{\beta \rightarrow \infty} \varepsilon_{\mathbf{F}}^{\beta, R}(u)$ exists but is not, in general, equal to $\lim_{\beta \rightarrow \infty} \varepsilon_{\mathbf{F}}^{\beta, R}$.

- (3) In the case of a insulator, the limit $\lim_{R \rightarrow \infty} \lim_{\beta \rightarrow \infty} \varepsilon_{\mathbf{F}}^{\beta, R}(u)$ exists but, in general, is not equal to $\varepsilon_{\mathbf{F}}^{\infty, \#}$.

- (4) In the case of a metal, the zero temperature and thermodynamic limits of the Fermi level exist and commute:

$$\lim_{R \rightarrow \infty} \lim_{\beta \rightarrow \infty} \varepsilon_{\mathbf{F}}^{\beta, R}(u) = \lim_{\beta \rightarrow \infty} \lim_{R \rightarrow \infty} \varepsilon_{\mathbf{F}}^{\beta, R}(u) = \lim_{\beta, R \rightarrow \infty} \varepsilon_{\mathbf{F}}^{\beta, R}.$$

Remark 4. It is important to note that we do not prove (3) in full generality. As we will see, perturbing the system may introduce eigenvalues into the band gap. We only prove (3) for the case where these eigenvalues are isolated. If the eigenvalues cluster as $R \rightarrow \infty$, or we have repeated eigenvalues, more arguments are needed.

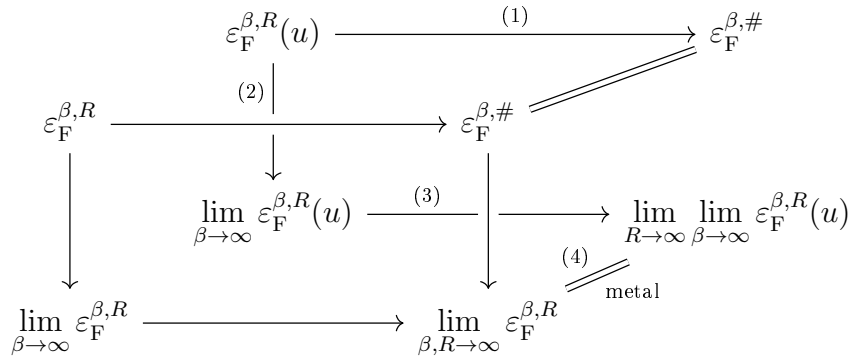


Figure 3.1: Commutative diagram to summarise both Theorem 1 and 2. Here, the horizontal arrows represent thermodynamic limits ($R \rightarrow \infty$) and vertical arrows are zero temperature limits ($\beta \rightarrow \infty$). The labels correspond to the numbering of the statements in Theorem 2.

We now informally discuss the main ideas in the proof of Theorem 2 and give some examples. A full proof is given in Chapter 4.

First of all, under the regularity assumptions on the functions $h^{\alpha\beta}$, we can use Taylor's Theorem to write

$$\mathcal{H}(x + u) = \mathcal{H} + P(u) \quad (3.1)$$

for some perturbation matrix $P(u)$. We will see that, under the assumption on the displacement, we will be able to write

$$P(u) = P^{(1)}(u) + P^{(2)}(u)$$

where $P^{(1)}(u)$ has finite (R -independent) rank and $P^{(2)}(u)$ is a small perturbation in the Frobenius matrix norm. In Proposition 3 and 4 below, we shall see how to deal with both of these terms.

First we look at an example of how perturbing the homogeneous Hamiltonian by a finite rank matrix affects the spectrum.

Example 4. *We shall consider the case of a one dimensional lattice with two atomic orbitals per atom and the homogeneous Hamiltonian given by a nearest neighbour model. That is, for atoms $\ell, k \in \Lambda_R$ and atomic orbitals indexed by $\alpha, \beta \in \{1, 2\}$, we have*

$$(\mathcal{H}_{hom}^R)_{\ell k}^{\alpha\beta} = \begin{cases} h_0^\alpha & \text{if } \ell = k \text{ and } \alpha = \beta \\ h_1^\alpha & \text{if } \ell \sim_{\Lambda_R} k \text{ and } \alpha = \beta \\ h_1^{12} & \text{if } \ell \sim_{\Lambda_R} k \text{ and } \alpha \neq \beta \\ 0 & \text{otherwise} \end{cases}$$

for some constants $h_0^1, h_0^2, h_1^1, h_1^{12}, h_1^2$. In Example 3, we gave a formula for the eigenvalues of this Hamiltonian and plotted the energy bands for some choice of constants that gave rise to a band gap (shown in Figure 2.2). We will now add a defect into the system and see how this affects the spectrum. We will position the defect at $0 \in \Lambda_R$ and suppose it has the form of a simple nearest neighbour interaction. That is, we define

$$(\mathcal{H}_{def}^R)_{\ell k}^{\alpha\beta} = \begin{cases} d_0^\alpha & \text{if } \ell = k = 0 \text{ and } \alpha = \beta \\ d_1^\alpha & \text{if } \ell, k \in \{0, \pm 1\}, \ell \sim_{\Lambda_R} k \text{ and } \alpha = \beta \\ d_1^{12} & \text{if } \ell, k \in \{0, \pm 1\}, \ell \sim_{\Lambda_R} k \text{ and } \alpha \neq \beta \\ 0 & \text{otherwise} \end{cases}$$

for some choice of constants $d_0^1, d_0^2, d_1^1, d_1^2, d_1^{12}$. The spectrum of $\mathcal{H}^R := \mathcal{H}_{hom}^R + \mathcal{H}_{def}^R$ is plotted in Figure 3.2.

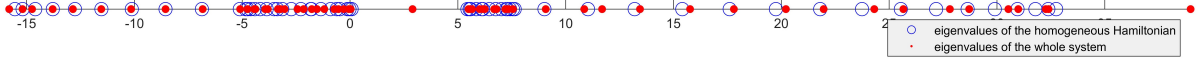


Figure 3.2: Spectrum of the homogeneous Hamiltonian alongside the spectrum of the system after adding in a defect at $\ell = 0 \in \Lambda$. Here, we take the simple nearest neighbour model with constants $h_0^1 = 0, h_0^2 = 10, h_1^1 = -1, h_1^2 = 10, h_1^{12} = 5, d_0^1 = d_0^2 = 10, d_1^1 = d_1^2 = 4, d_1^{12} = -1$ and system size $R = 30$. The spectrum of the homogeneous system is shown in blue circles and that of the defective system is in red.

We can see that the eigenvalues of the defective system are positioned between the eigenvalues of the homogeneous Hamiltonian. This interlacing property is the content of Proposition 3 stated below.

For a real symmetric $n \times n$ matrix M , we use $\lambda_j(M)$ to denote the j^{th} smallest eigenvalue of M . Further, we set $\lambda_j(\cdot) = -\infty$ for $j \leq 0$ and $\lambda_j(\cdot) = +\infty$ for $j > n$. As alluded to in Example 4, we will show that the eigenvalues of a finite rank updated matrix satisfy the following interlacing property:

Proposition 3 (Interlacing Property). *Suppose we have $n \times n$ real symmetric matrices A and B and further suppose that B has rank k . Then, there exists a k' that does not depend upon n such that,*

$$\lambda_{j-k'}(A) \leq \lambda_j(A + B) \leq \lambda_{j+k'}(A) \quad \text{for all } 1 \leq j \leq n.$$

By Proposition 3, the finite rank perturbation introduces only a finite (R -independent) number of eigenvalues outside the homogeneous density of states. Therefore, after only perturbing the system by a finite rank matrix, the density of states (in the thermodynamic limit) will be unchanged. This means that the thermodynamic limit of the defective Fermi level is the same as for the homogeneous case.

Despite this, in the case of an insulator, the newly introduced eigenvalues in the band gap have an impact on the zero temperature limit of the Fermi level corresponding to a finite domain. We now see that this is the case for Example 4.

Example 4 (continued). *Figure 3.3 displays the zero temperature limits of the homogeneous and defective Fermi levels from Example 4 for a number of system sizes. It should be clear that these are different for each fixed R and that they converge to different values in the thermodynamic limit.*

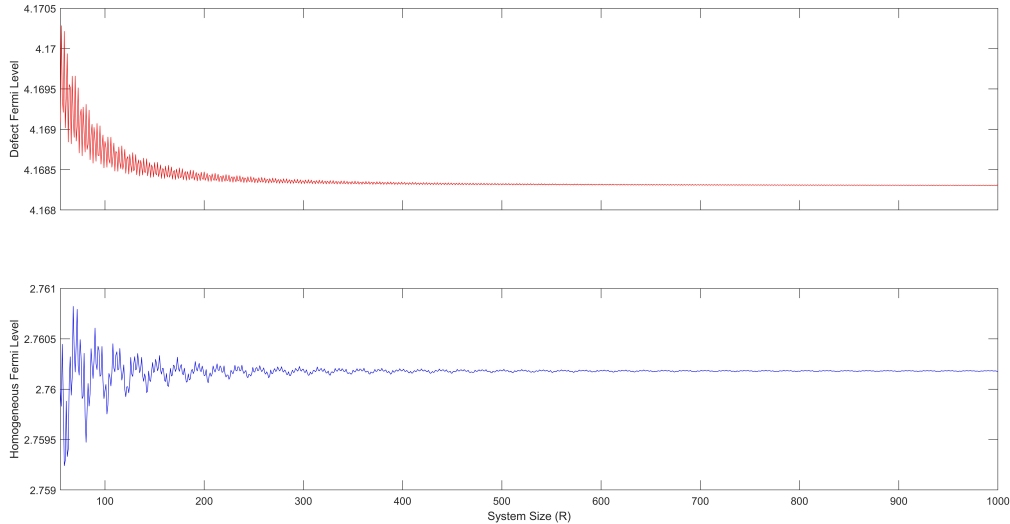


Figure 3.3: Zero temperature limits of the Fermi levels for increasing system size in the case of a defect (red) and in the case of the homogeneous Hamiltonian (blue). Again, we take the simple nearest neighbour model from Example 4 with constants $h_0^1 = 0, h_0^2 = 10, h_1^1 = -1, h_1^2 = 10, h_1^{12} = 5, d_0^1 = d_0^2 = 10, d_1^1 = d_1^2 = 4, d_1^{12} = -1$ and system size R between 50 and 1000.

We have only discussed the case of a finite rank perturbation but, as we will see in Chapter 4, unless the displacement is compactly supported, the perturbation in (3.1) will not have finite R -independent rank. Instead, we will be able to write it as the sum of a finite rank matrix and a matrix that is small in the sense of the Frobenius norm. We will then apply Weyl's Perturbation Theorem:

Proposition 4 (Weyl). *Suppose we have a real symmetric matrix A and perturb it by another real symmetric matrix P . Then, for each $1 \leq k \leq n$, we have that*

$$|\lambda_k(A + P) - \lambda_k(A)| \leq \|P\|$$

for any sub-multiplicative matrix norm $\|\cdot\|$.

A proof of this result along with more general inequalities involving the eigenvalues of perturbed matrices can be found in [1].

In our case, this means that adding on the small perturbation does not change the spectrum by much. Therefore, we will be able to show that the thermodynamic limit of the Fermi level at finite temperature coincides with the non-defect case. Further, we shall prove that, in the case of a metal, the thermodynamic and zero temperature limits of the Fermi level exist and commute.

Chapter 4

Proofs

Proof of Proposition 1. Let us fix $R \in \mathbb{N}$ and $k \in \text{BZ}_R$. We set $\psi = (c^\alpha(k)e^{i\ell \cdot k})_{\ell \in \Lambda}^{\alpha \in \Xi}$ for some constants $c^\alpha(k)$, normalised so that $\sum_\alpha |c^\alpha(k)|^2 = 1$. Now we can see that for $\alpha \in \Xi$ and $\ell \in \Lambda_R$ we have

$$\begin{aligned}
(\mathcal{H}^R \psi)_\ell^\alpha &= \sum_{\beta \in \Xi} \sum_{m \in \Lambda_R} \mathcal{H}_{\ell m}^{\alpha\beta} \psi^\beta(m) \\
&= \sum_{\beta \in \Xi} \sum_{m \in \Lambda_R} \mathcal{H}_{\ell m}^{\alpha\beta} c^\beta(k) e^{im \cdot k} \\
&= \sum_{\beta \in \Xi} \sum_{m \in \Lambda_R} h^{\alpha\beta}(m - \ell) e^{i(m-\ell) \cdot k} c^\beta(k) e^{i\ell \cdot k} && \text{(by definition of } h^{\alpha\beta}) \\
&= \sum_{\beta \in \Xi} \left(\sum_{m \in \Lambda_R} h^{\alpha\beta}(m) e^{im \cdot k} \right) c^\beta(k) e^{i\ell \cdot k} && \text{(by symmetry \& choice of } k) \\
&= \sum_{\beta \in \Xi} \widehat{h^{\alpha\beta}}(k) \psi^\beta(\ell)
\end{aligned}$$

Now, ψ is an eigenvector if and only if there exists some associated eigenvalue $\varepsilon(k)$ satisfying the following equation

$$\det(\hat{h}(k) - \varepsilon(k)I_{n_\Xi}) = 0. \quad (4.1)$$

This means, for each fixed $k \in \text{BZ}_R$, the n_Ξ solutions (counting multiplicities) of this equation are eigenvalues of $\mathcal{H}^R$. Since $\hat{h}$ is a continuous function on $\overline{\text{BZ}}$, the coefficients of the polynomial equation determining eigenvalues are continuous functions of k as well. The complex roots of this equation, which we will denote by $\{\varepsilon^\alpha(k)\}_{\alpha \in \Xi}$, are thus also continuous functions of k [10]. $\square$

4.1 Proof of Theorem 1

In the following, we shall use the notation

$$\sigma(\mathcal{H}^R) = \{\varepsilon_1^R \leq \dots \leq \varepsilon_{\#\sigma_R}^R\}$$

where $\#\sigma_R := \#\sigma(\mathcal{H}^R)$ of course depends on R but also on the number of orbitals that we are considering. Moreover, we define $\underline{\varepsilon}_R := \varepsilon_{\frac{\#\sigma_R}{2}}^R$ and $\bar{\varepsilon}_R := \varepsilon_{\frac{\#\sigma_R}{2}+1}^R$. Also, recall that we use $[\underline{\varepsilon}, \bar{\varepsilon}]$ to denote the band gap of the spectrum in the thermodynamic limit. That is, $[\underline{\varepsilon}, \bar{\varepsilon}]$ is the maximal interval such that $\rho([\underline{\varepsilon}, \bar{\varepsilon}]) = 0$ and $\rho((-\infty, \underline{\varepsilon})) = \rho((\bar{\varepsilon}, \infty)) = \frac{1}{2}$.

We first prove an important lemma that will be used a number of times in the proof of Theorem 1:

Lemma 1. *The local particle number satisfies the uniform convergence properties:*

- (1) $\lim_{R \rightarrow \infty} \|N_{\#}^{\beta, R}(\cdot) - N_{\#}^{\beta, \#}(\cdot)\|_{L^\infty} = 0$ for each finite $\beta > 0$,
- (2) $\lim_{\beta, R \rightarrow \infty} \|N_{\#}^{\beta, R}(\cdot) - N_{\#}^{\infty, \#}(\cdot)\|_{L^\infty} = 0$.

In particular, property (2) implies that

$$\begin{aligned} N_{\#}^{\beta, \#}(\cdot) &\rightarrow N_{\#}^{\infty, \#}(\cdot) \quad \text{uniformly as } \beta \rightarrow \infty \text{ and} \\ N_{\#}^{\infty, R}(\cdot) &\rightarrow N_{\#}^{\infty, \#}(\cdot) \quad \text{uniformly as } R \rightarrow \infty. \end{aligned}$$

For all finite $R > 0$ and compact intervals K such that $N_{\#}^{\infty, R}(\cdot)$ is a constant function on K , we have

$$(3) \quad \lim_{\beta \rightarrow \infty} \|N_{\#}^{\beta, R}(\cdot) - N_{\#}^{\infty, R}(\cdot)\|_{L^\infty(K)} = 0.$$

Proof of (1). Fix $\beta \in (0, \infty)$. We know that $N_{\#}^{\beta, R}(\cdot)$ is a sequence of differentiable functions converging pointwise to $N_{\#}^{\beta, \#}(\cdot)$ as $R \rightarrow \infty$. Furthermore,

$$\left| \frac{\partial}{\partial \tau} N_{\#}^{\beta, R}(\tau) \right| = \frac{2}{n_{\Xi} \#\text{BZ}_R} \sum_{\alpha \in \Xi} \sum_{k \in \text{BZ}_R} -\beta f'_{\beta}(\varepsilon^{\alpha}(k) - \tau) \leq \frac{1}{2}\beta$$

and so $\frac{\partial}{\partial \tau} N_{\#}^{\beta, R}(\cdot)$ is uniformly bounded in R (and τ). Now, a consequence of the Arzelà–Ascoli Theorem [18], implies the uniform convergence of $N_{\#}^{\beta, R}(\cdot)$ as $R \rightarrow \infty$. $\square$

Proof of (2). For now, we assume the following properties which are shown after we have seen the main ideas in the proof.

- (a) $\rho([\tau - \delta, \tau + \delta]) \rightarrow 0$ uniformly in τ as $\delta \rightarrow 0$,
- (b) For all fixed $\delta > 0$, $\rho_R([\tau - \delta, \tau + \delta]) \rightarrow \rho([\tau - \delta, \tau + \delta])$ uniformly in τ as $R \rightarrow \infty$,
- (c) $\rho_R((-\infty, \tau)) \rightarrow \rho((-\infty, \tau))$ uniformly in τ as $R \rightarrow \infty$.

Let us fix $\eta > 0$. By (a), we choose $\delta > 0$ such that $\rho([\tau - \delta, \tau + \delta]) \leq \frac{\eta}{6}$ for all τ . Next, by (b) and (c), we can choose $R' > 0$ sufficiently large such that $\rho_R([\tau - \delta, \tau + \delta]) \leq \frac{\eta}{3}$ and $|(\rho_R - \rho)((-\infty, \tau))| \leq \frac{\eta}{6}$ for all $R > R'$ and all τ . Choosing β' sufficiently large such that $f_\beta(\delta) \leq \frac{\eta}{6}$ for all $\beta > \beta'$ allows to conclude: for all $\tau \in \mathbb{R}$,

$$\begin{aligned}
|N_{\#}^{\beta, R}(\tau) - N_{\#}^{\infty, \#}(\tau)| &\leq |N_{\#}^{\beta, R}(\tau) - N_{\#}^{\infty, R}(\tau)| + |N_{\#}^{\infty, R}(\tau) - N_{\#}^{\infty, \#}(\tau)| \\
&= 2 \left| \int (f_\beta - f_\infty)(\cdot - \tau) d\rho_R \right| + 2 \left| \int f_\infty(\cdot - \tau) d(\rho_R - \rho) \right| \\
&\leq \rho_R([\tau - \delta, \tau + \delta]) + 2f_\beta(\delta) + 2|(\rho_R - \rho)((-\infty, \tau))| \leq \eta
\end{aligned}$$

for all $\beta > \beta'$ and $R > R'$.

Now we return to the proofs of properties (a), (b) and (c):

(a) By **(E)**, it is clear that, for each τ , we have

$$\begin{aligned}
\lim_{\delta \rightarrow 0} \rho([\tau - \delta, \tau + \delta]) &= \frac{1}{n_{\Xi}} \sum_{\alpha \in \Xi} \frac{1}{|\text{BZ}|} \lim_{\delta \rightarrow 0} |\{k \in \text{BZ} : \varepsilon^\alpha(k) \in [\tau - \delta, \tau + \delta]\}| \\
&= \frac{1}{n_{\Xi}} \sum_{\alpha \in \Xi} \frac{1}{|\text{BZ}|} \left| \bigcap_{\delta > 0} (\varepsilon^\alpha)^{-1}([\tau - \delta, \tau + \delta]) \right| \\
&= \frac{1}{n_{\Xi}} \sum_{\alpha \in \Xi} \frac{1}{|\text{BZ}|} |(\varepsilon^\alpha)^{-1}(\{\tau\})| = 0.
\end{aligned}$$

Now, since the sequence $\{\rho([\tau - \delta, \tau + \delta])\}_\delta$ is monotonically decreasing (as $\delta \rightarrow 0$) for all τ , we can invoke Dini's Theorem [12, Theorem 12.1] to conclude that the convergence is in fact uniform in τ .

(b), (c): We shall apply the following proposition:

Proposition 5 (Cited in [2]). *Let $\mathbb{P}$ and $\mathbb{P}_n$ be probability measures on $\mathbb{R}$. Suppose*

1. *There is a dense set D such that $\mathbb{P}_n((-\infty, t]) \rightarrow \mathbb{P}((-\infty, t])$ as $n \rightarrow \infty$ for all $t \in D$,*
2. *$t \mapsto \mathbb{P}((-\infty, t])$ is a continuous function on $\mathbb{R}$.*

Then we necessarily have that

$$\sup_{t \in \mathbb{R}} |\mathbb{P}_n((-\infty, t]) - \mathbb{P}((-\infty, t])| \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

It is clear that the conditions from Proposition 5 are satisfied by $D = \mathbb{R}$ and the measures $\mathbb{P}_n = \rho_n$ and $\mathbb{P} = \rho$. This means that we have

$$\sup_{\tau \in \mathbb{R}} |\rho_R((-\infty, \tau]) - \rho((-\infty, \tau])| \rightarrow 0 \quad \text{as } R \rightarrow \infty.$$

By **(E)**, $\rho_R(\{\tau\}) \rightarrow 0$ uniformly in τ and so we have (c). By noting that $[\tau - \delta, \tau + \delta] = (-\infty, \tau + \delta] \setminus (-\infty, \tau - \delta)$ we can conclude that (b) also holds. $\square$

Proof of (3). Fix a compact K such that $N_{\#}^{\infty, R}|_K$ is constant. We know that $N^{\beta, R}(\cdot)$ is a sequence of differentiable functions on K converging pointwise to the continuous function $N_{\#}^{\infty, R}|_K(\cdot)$ as $\beta \rightarrow \infty$. Furthermore, since there is no eigenvalue of $\mathcal{H}^R$ contained in K , we have

$$\begin{aligned} \left| \frac{\partial}{\partial \tau} N_{\#}^{\beta, R}(\tau) \right| &= \frac{2}{n_{\Xi}} \frac{1}{\#\text{BZ}_R} \sum_{\alpha \in \Xi} \sum_{k \in \text{BZ}_R} -\beta f'_{\beta}(\varepsilon^{\alpha}(k) - \tau) \\ &\leq 2\beta \max_{\varepsilon \in \sigma(\mathcal{H}^R)} \max_{E \in K} f_{\beta}(|\varepsilon - E|) \rightarrow 0 \end{aligned}$$

as $\beta \rightarrow \infty$. In particular, $\frac{\partial}{\partial \tau} N_{\#}^{\beta, R}(\cdot)$ is uniformly bounded in β (and τ). Again, we apply a consequence of the Arzelà–Ascoli Theorem [18], to conclude. $\square$

We will now break the proof of Theorem 1 into five steps that correspond to each of the limits in the commutative diagram.

Step 1. *Using the assumptions from Theorem 1, we have*

$$\lim_{R \rightarrow \infty} \varepsilon_{\text{F}}^{\beta, R} = \varepsilon_{\text{F}}^{\beta, \#} \quad \text{for each finite } \beta > 0.$$

$$\begin{array}{ccc} \varepsilon_{\text{F}}^{\beta, R} & \xrightarrow{(1)} & \varepsilon_{\text{F}}^{\beta, \#} \\ (3) \downarrow & \text{~~~~~} (5) \text{~~~~~} & \downarrow (2) \\ \frac{1}{2}(\underline{\varepsilon}_R + \bar{\varepsilon}_R) & \xrightarrow{(4)} & \frac{1}{2}(\underline{\varepsilon} + \bar{\varepsilon}) \end{array}$$

Proof. Fix $\beta \in (0, \infty)$. By property (1) of Lemma 1 and the fact that $N_{\#}^{\beta, R}(\varepsilon_{\text{F}}^{\beta, R}) = N_{\#}^{\beta, \#}(\varepsilon_{\text{F}}^{\beta, \#}) = 1$, we have

$$\begin{aligned} |N_{\#}^{\beta, \#}(\varepsilon_{\text{F}}^{\beta, R}) - N_{\#}^{\beta, \#}(\varepsilon_{\text{F}}^{\beta, \#})| &\leq |N_{\#}^{\beta, \#}(\varepsilon_{\text{F}}^{\beta, R}) - N_{\#}^{\beta, R}(\varepsilon_{\text{F}}^{\beta, R})| + |N_{\#}^{\beta, R}(\varepsilon_{\text{F}}^{\beta, R}) - N_{\#}^{\beta, \#}(\varepsilon_{\text{F}}^{\beta, \#})| \\ &\longrightarrow 0 \quad \text{as } R \rightarrow \infty. \end{aligned}$$

Since, $N_{\#}^{\beta, \#}(\cdot)$ is a strictly increasing continuous function, there exists a continuous inverse. In particular, we have $\varepsilon_{\text{F}}^{\beta, R} \rightarrow \varepsilon_{\text{F}}^{\beta, \#}$ as $R \rightarrow \infty$. $\square$

Step 2. Using the assumptions from Theorem 1, we have

$$\lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta, \#} = \varepsilon_F^{\infty, \#} := \frac{1}{2}(\underline{\varepsilon} + \bar{\varepsilon}).$$

$$\begin{array}{ccc} \varepsilon_F^{\beta, R} & \xrightarrow{(1)} & \varepsilon_F^{\beta, \#} \\ (3) \downarrow & \text{~~~~~} (5) \text{~~~~~} & \downarrow (2) \\ \frac{1}{2}(\underline{\varepsilon}_R + \bar{\varepsilon}_R) & \xrightarrow{(4)} & \frac{1}{2}(\underline{\varepsilon} + \bar{\varepsilon}) \end{array}$$

Proof. We first observe that $\varepsilon_F^{\beta, R} \in [\underline{\sigma}, \bar{\sigma}]$ for each β, R . Indeed, we know that the spectrum of the Hamiltonian is R -uniformly bounded in $[\underline{\sigma}, \bar{\sigma}]$ and thus $N_{\#}^{\beta, R}(\tau) < 1$ for all $\tau < \underline{\sigma}$ and similarly $N_{\#}^{\beta, R}(\tau) > 1$ for all $\tau > \bar{\sigma}$. Furthermore, by Step 1, we have $\varepsilon_F^{\beta, \#} \in [\underline{\sigma}, \bar{\sigma}]$. Thus, there exists a convergent subsequence (which we do not relabel) such that $\varepsilon_F^{\beta, \#} \rightarrow \varepsilon$ as $\beta \rightarrow \infty$ for some $\varepsilon \in [\underline{\sigma}, \bar{\sigma}]$. We will first show that $N_{\#}^{\infty, \#}(\varepsilon) = 1$ and so $\varepsilon \in [\underline{\sigma}, \bar{\sigma}]$.

By the continuity of $N_{\#}^{\infty, \#}(\cdot)$ and uniform convergence of $\{N_{\#}^{\beta, \#}(\cdot)\}_{\beta}$, we can see that $N_{\#}^{\infty, \#}(\varepsilon) = 1$. Indeed:

$$\begin{aligned} |N_{\#}^{\infty, \#}(\varepsilon) - 1| &\leq |N_{\#}^{\infty, \#}(\varepsilon) - N_{\#}^{\infty, \#}(\varepsilon_F^{\beta, \#})| + |N_{\#}^{\infty, \#}(\varepsilon_F^{\beta, \#}) - N_{\#}^{\beta, \#}(\varepsilon_F^{\beta, \#})| \\ &\longrightarrow 0 \quad \text{as } \beta \rightarrow \infty. \end{aligned}$$

Therefore, in the case of a metal, $\lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta, \#} = \varepsilon_F^{\infty, \#}$.

On the other hand, in the case of an insulator, we have shown that if $\varepsilon_F^{\beta, \#} \rightarrow \varepsilon$ along a subsequence, then we necessarily have $\varepsilon \in [\underline{\varepsilon}, \bar{\varepsilon}]$. We can use the fact that $f_{\beta}(E) + f_{\beta}(-E) = 1$ to obtain

$$1 = N_{\#}^{\beta, \#}(\varepsilon_F^{\beta, \#}) = 2 \left(\int_{\{\tau < \varepsilon_F^{\beta, \#}\}} \{1 - f_{\beta}(|\tau - \varepsilon_F^{\beta, \#}|)\} d\rho(\tau) + \int_{\{\tau \geq \varepsilon_F^{\beta, \#}\}} f_{\beta}(|\tau - \varepsilon_F^{\beta, \#}|) d\rho(\tau) \right).$$

Since

$$N_{\#}^{\infty, \#}(\varepsilon_F^{\beta, \#}) = 2\rho((-\infty, \varepsilon_F^{\beta, \#})) \xrightarrow{\beta \rightarrow \infty} N_{\#}^{\infty, \#}(\varepsilon) = 1,$$

for each fixed $\delta > 0$, $|\rho((-\infty, \varepsilon_F^{\beta, \#})) - \frac{1}{2}| < \delta$ for sufficiently large β . Thus, we have

$$\left| \int_{\{\tau < \varepsilon_F^{\beta, \#}\}} f_{\beta}(|\tau - \varepsilon_F^{\beta, \#}|) d\rho(\tau) - \int_{\{\tau \geq \varepsilon_F^{\beta, \#}\}} f_{\beta}(|\tau - \varepsilon_F^{\beta, \#}|) d\rho(\tau) \right| \leq \delta \quad (4.2)$$

for all sufficiently large β . Therefore, for each $r > 0$ and sufficiently large β , we obtain

$$\begin{aligned} \rho([\varepsilon_F^{\beta, \#} - r, \varepsilon_F^{\beta, \#}]) f_{\beta}(r) &\leq \int_{\{\tau < \varepsilon_F^{\beta, \#}\}} f_{\beta}(|\tau - \varepsilon_F^{\beta, \#}|) d\rho(\tau) \\ &\leq \int_{\{\tau \geq \varepsilon_F^{\beta, \#}\}} f_{\beta}(|\tau - \varepsilon_F^{\beta, \#}|) d\rho(\tau) + \delta \\ &\leq \left(\frac{1}{2} + \delta \right) f_{\beta}(|\bar{\varepsilon} - \varepsilon_F^{\beta, \#}|) + \delta \end{aligned}$$

and similarly,

$$\rho((\varepsilon_F^{\beta,\#}, \varepsilon_F^{\beta,\#} + r]) f_\beta(r) \leq \left(\frac{1}{2} + \delta\right) f_\beta(|\underline{\varepsilon} - \varepsilon_F^{\beta,\#}|) + \delta.$$

Now, since $\delta > 0$ was arbitrary, we necessarily have that

$$\frac{\rho([\varepsilon_F^{\beta,\#} - r, \varepsilon_F^{\beta,\#})) f_\beta(r)}{f_\beta(|\bar{\varepsilon} - \varepsilon_F^{\beta,\#}|)} \quad \text{and} \quad \frac{\rho((\varepsilon_F^{\beta,\#}, \varepsilon_F^{\beta,\#} + r]) f_\beta(r)}{f_\beta(|\underline{\varepsilon} - \varepsilon_F^{\beta,\#}|)} \quad (4.3)$$

are bounded by some constant, C , independent of both r and β . We now fix $\delta > 0$ sufficiently small such that $\underline{\varepsilon} + \delta < \bar{\varepsilon} - \delta$ and note that $\varepsilon_F^{\beta,\#}$ is contained within $(\varepsilon - \delta, \varepsilon + \delta)$ for sufficiently large β .

At this point, we suppose that $\varepsilon \leq \frac{1}{2}(\underline{\varepsilon} + \bar{\varepsilon})$. For (4.3) to hold, we require

$$\frac{\rho([\varepsilon + \delta - r, \varepsilon - \delta]) f_\beta(r)}{f_\beta(|\bar{\varepsilon} - (\varepsilon + \delta)|)} \leq \frac{\rho([\varepsilon_F^{\beta,\#} - r, \varepsilon_F^{\beta,\#})) f_\beta(r)}{f_\beta(|\bar{\varepsilon} - \varepsilon_F^{\beta,\#}|)} \leq C.$$

We have used $\varepsilon \leq \frac{1}{2}(\underline{\varepsilon} + \bar{\varepsilon})$ to conclude that $f_\beta(|\bar{\varepsilon} - \varepsilon_F^{\beta,\#}|) \leq f_\beta(|\bar{\varepsilon} - (\varepsilon + \delta)|)$. Choosing

$$r > |\underline{\varepsilon} - (\varepsilon + \delta)| \quad (4.4)$$

ensures $\varepsilon + \delta - r < \underline{\varepsilon}$ and so we have $\rho([\varepsilon + \delta - r, \varepsilon - \delta]) > 0$. Now, we simply need

$$\frac{f_\beta(r)}{f_\beta(|\bar{\varepsilon} - (\varepsilon + \delta)|)} \leq C$$

where C is independent of β . Since, for $E_1, E_2 > 0$, we have

$$\lim_{\beta \rightarrow \infty} \frac{f_\beta(E_1)}{f_\beta(E_2)} = \begin{cases} 0 & \text{if } E_1 > E_2 \\ 1 & \text{if } E_1 = E_2 \\ \infty & \text{if } E_1 < E_2, \end{cases}$$

we require $r \geq |\bar{\varepsilon} - (\varepsilon + \delta)|$. We have shown that if $r > 0$ satisfies (4.4) then $r \geq \bar{\varepsilon} - (\varepsilon + \delta)$.

Therefore, we necessarily have $\bar{\varepsilon} - (\varepsilon + \delta) \leq (\varepsilon + \delta) - \underline{\varepsilon}$ and thus $\varepsilon \geq \frac{1}{2}(\underline{\varepsilon} + \bar{\varepsilon}) - \delta$.

In the exact same way, if we instead assume $\varepsilon \geq \frac{1}{2}(\underline{\varepsilon} + \bar{\varepsilon})$ and use the uniform boundedness

$$\frac{\rho([\varepsilon + \delta, \varepsilon - \delta + r]) f_\beta(r)}{f_\beta(|\underline{\varepsilon} - \varepsilon_F^{\beta,\#}|)} \leq \frac{\rho((\varepsilon_F^{\beta,\#}, \varepsilon_F^{\beta,\#} + r]) f_\beta(r)}{f_\beta(|\underline{\varepsilon} - \varepsilon_F^{\beta,\#}|)} \leq C$$

we can conclude that $\varepsilon \leq \frac{1}{2}(\underline{\varepsilon} + \bar{\varepsilon}) + \delta$. Since the choice of $\delta > 0$ was arbitrary, we have shown that $\varepsilon = \frac{1}{2}(\underline{\varepsilon} + \bar{\varepsilon})$. $\square$

Step 3. Using the assumptions from Theorem 1, we have

$$\lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta, R} = \frac{1}{2}(\underline{\varepsilon}_R + \bar{\varepsilon}_R).$$

$$\begin{array}{ccc} \varepsilon_F^{\beta, R} & \xrightarrow{(1)} & \varepsilon_F^{\beta, \#} \\ (3) \downarrow & \text{wavy} (5) \text{ wavy} & \downarrow (2) \\ \frac{1}{2}(\underline{\varepsilon}_R + \bar{\varepsilon}_R) & \xrightarrow{(4)} & \frac{1}{2}(\underline{\varepsilon} + \bar{\varepsilon}) \end{array}$$

Proof. For simplicity of notation, we define

$$\underline{m} := \# \left\{ j \leq \frac{\#\sigma_R}{2} : \varepsilon_j^R = \underline{\varepsilon}_R \right\} \quad \text{and} \quad \bar{m} := \# \left\{ j \geq \frac{\#\sigma_R}{2} + 1 : \varepsilon_j^R = \bar{\varepsilon}_R \right\}.$$

In the case that $\underline{\varepsilon}_R \neq \bar{\varepsilon}_R$, $\underline{m}, \bar{m}$ are simply the multiplicities of $\underline{\varepsilon}_R, \bar{\varepsilon}_R$ respectively.

Step 3(a): We will first show that, if the Fermi levels converge along a subsequence to some value, then this value must be inside $[\underline{\varepsilon}_R, \bar{\varepsilon}_R]$.

Note that, in the case that $\underline{\varepsilon}_R \neq \bar{\varepsilon}_R$, we have

$$\begin{aligned} N_{\#}^{\infty, R}(\underline{\varepsilon}_R) &= 1 - \frac{\underline{m}}{\#\sigma_R}, \quad N_{\#}^{\infty, R}(\bar{\varepsilon}_R) = 1 + \frac{\bar{m}}{\#\sigma_R}, \\ N_{\#}^{\infty, R}(\tau) &= 1 \quad \text{for } \tau \in (\underline{\varepsilon}_R, \bar{\varepsilon}_R) \end{aligned} \tag{4.5}$$

and for $\underline{\varepsilon}_R = \bar{\varepsilon}_R$, we have

$$\begin{aligned} \lim_{\tau \uparrow \underline{\varepsilon}_R} N_{\#}^{\infty, R}(\tau) &= 1 - \frac{2\underline{m}}{\#\sigma_R}, \\ N_{\#}^{\infty, R}(\underline{\varepsilon}_R = \bar{\varepsilon}_R) &= 1 + \frac{\bar{m} - \underline{m}}{\#\sigma_R}, \\ \lim_{\tau \downarrow \bar{\varepsilon}_R} N_{\#}^{\infty, R}(\tau) &= 1 + \frac{2\bar{m}}{\#\sigma_R}. \end{aligned} \tag{4.6}$$

After possibly taking a subsequence (which we do not relabel), we can find an $\varepsilon \in [\underline{\varepsilon}, \bar{\varepsilon}]$ such that $\varepsilon_F^{\beta, R} \rightarrow \varepsilon$ as $\beta \rightarrow \infty$.

First, let us suppose that $\varepsilon \notin \sigma(\mathcal{H}^R)$ and so we can find adjacent eigenvalues $\varepsilon^{(1)} < \varepsilon^{(2)}$ such that $\varepsilon \in (\varepsilon^{(1)}, \varepsilon^{(2)})$. Since, $N_{\#}^{\infty, R}(\cdot)$ is constant on the interval $(\varepsilon^{(1)}, \varepsilon^{(2)})$, we know, by property (3) of Lemma 1, that $N_{\#}^{\beta, R}(\cdot) \rightarrow N_{\#}^{\infty, R}(\cdot)$ uniformly on compact subsets $K \subset (\varepsilon^{(1)}, \varepsilon^{(2)})$. Since, for large enough β , the sequence $\{\varepsilon_F^{\beta, R}\}$ is contained in such a compact set, we have

$$\begin{aligned} |N_{\#}^{\infty, R}(\varepsilon) - 1| &\leq |N_{\#}^{\infty, R}(\varepsilon) - N_{\#}^{\infty, R}(\varepsilon_F^{\beta, R})| + |N_{\#}^{\infty, R}(\varepsilon_F^{\beta, R}) - N_{\#}^{\beta, R}(\varepsilon_F^{\beta, R})| \\ &\rightarrow 0 \text{ as } \beta \rightarrow \infty. \end{aligned}$$

Therefore, we can conclude that $N_{\#}^{\infty, R}(\varepsilon) = 1$. Using (4.5) and (4.6), it is clear that $\underline{\varepsilon}_R \neq \bar{\varepsilon}_R$ and $\varepsilon \in (\underline{\varepsilon}_R, \bar{\varepsilon}_R)$.

Now we instead assume that ε is an eigenvalue and find

$$\varepsilon^{(1)} := \max\{\tau \in \sigma(\mathcal{H}^R) : \tau < \varepsilon\} \quad \text{and} \quad \varepsilon^{(2)} := \min\{\tau \in \sigma(\mathcal{H}^R) : \tau > \varepsilon\}.$$

Let us choose $\delta > 0$ sufficiently small such that $\varepsilon^{(1)} < \varepsilon - \delta < \varepsilon < \varepsilon + \delta < \varepsilon^{(2)}$ and find a $\beta' > 0$ large enough so that $\varepsilon_F^{\beta,R} \in (\varepsilon - \delta, \varepsilon + \delta)$ for all $\beta > \beta'$. As $N_{\#}^{\beta,R}(\cdot)$ is strictly increasing for each $\beta > 0$, we know that

$$N_{\#}^{\beta,R}(\varepsilon - \delta) < N_{\#}^{\beta,R}(\varepsilon_F^{\beta,R}) < N_{\#}^{\beta,R}(\varepsilon + \delta) \quad \text{for all } \beta > \beta'.$$

We can first suppose that $\underline{\varepsilon}_R \neq \bar{\varepsilon}_R$. If $\varepsilon < \underline{\varepsilon}_R$ then, as $N_{\#}^{\beta,R}(\cdot)$ converges pointwise to $N_{\#}^{\infty,R}(\cdot)$ as $\beta \rightarrow \infty$, we have for all $\beta > \beta'$

$$1 = N_{\#}^{\beta,R}(\varepsilon_F^{\beta,R}) \leq N_{\#}^{\infty,R}(\varepsilon + \delta) < N_{\#}^{\infty,R}(\underline{\varepsilon}_R) = 1 - \frac{\underline{m}}{\#\sigma_R} < 1.$$

On the other hand, if $\varepsilon > \bar{\varepsilon}_R$ then for all $\beta > \beta'$

$$1 = N_{\#}^{\beta,R}(\varepsilon_F^{\beta,R}) \geq N_{\#}^{\infty,R}(\varepsilon - \delta) > N_{\#}^{\infty,R}(\bar{\varepsilon}_R) = 1 + \frac{\bar{m}}{\#\sigma_R} > 1.$$

In both these cases we have a contradiction and thus $\varepsilon \in [\underline{\varepsilon}_R, \bar{\varepsilon}_R]$ and so $\varepsilon = \underline{\varepsilon}_R$ or $\varepsilon = \bar{\varepsilon}_R$.

In the case that $\underline{\varepsilon}_R = \bar{\varepsilon}_R$ we can use a similar argument: if $\varepsilon < \underline{\varepsilon}_R = \bar{\varepsilon}_R$ then $\varepsilon + \delta < \underline{\varepsilon}_R$ and so for all $\beta > \beta'$

$$1 = N_{\#}^{\beta,R}(\varepsilon_F^{\beta,R}) \leq N_{\#}^{\infty,R}(\varepsilon + \delta) = 1 - \frac{2\underline{m}}{\#\sigma_R} < 1.$$

On the other hand, if $\varepsilon > \underline{\varepsilon}_R = \bar{\varepsilon}_R$ then $\varepsilon - \delta > \underline{\varepsilon}_R$ and so for all $\beta > \beta'$

$$1 = N_{\#}^{\beta,R}(\varepsilon_F^{\beta,R}) \geq N_{\#}^{\infty,R}(\varepsilon - \delta) = 1 + \frac{2\bar{m}}{\#\sigma_R} > 1.$$

In both these cases we have a contradiction and so $\varepsilon = \underline{\varepsilon}_R = \bar{\varepsilon}_R$. We can conclude from the above that if $\varepsilon_F^{\beta,R} \rightarrow \varepsilon$ along a subsequence, then we necessarily have that $\varepsilon \in [\underline{\varepsilon}_R, \bar{\varepsilon}_R]$.

Step 3(b): Now, we will show that either $\underline{\varepsilon}_R = \bar{\varepsilon}_R$ or the Fermi level is eventually contained in $(\underline{\varepsilon}_R, \bar{\varepsilon}_R)$ for large enough β . This will be important for the final step where we prove that $\varepsilon = \frac{1}{2}(\underline{\varepsilon}_R + \bar{\varepsilon}_R)$.

We assume that $\underline{\varepsilon}_R \neq \bar{\varepsilon}_R$. If the Fermi level is not eventually contained in $(\underline{\varepsilon}_R, \bar{\varepsilon}_R)$, there exists some subsequence (which we do not relabel) such that $\varepsilon_F^{\beta,R} \leq \underline{\varepsilon}_R$ for each β or such that $\varepsilon_F^{\beta,R} \geq \bar{\varepsilon}_R$ for each β . We only consider the former as the latter can be treated in the exact same way. This gives

$$1 = N_{\#}^{\beta,R}(\varepsilon_F^{\beta,R}) \leq N_{\#}^{\beta,R}(\underline{\varepsilon}_R) \rightarrow N_{\#}^{\infty,R}(\underline{\varepsilon}_R) = 1 - \frac{\underline{m}}{\#\sigma_R} < 1$$

as $\beta \rightarrow \infty$, a contradiction.

Step 3(c): In the case that $\underline{\varepsilon}_R = \bar{\varepsilon}_R$, we have already shown that $\varepsilon_F^{\beta,R} \rightarrow \frac{1}{2}(\underline{\varepsilon}_R + \bar{\varepsilon}_R)$ as $\beta \rightarrow \infty$. We will therefore assume that $\underline{\varepsilon}_R \neq \bar{\varepsilon}_R$. We have seen that $\varepsilon_F^{\beta,R} \in (\underline{\varepsilon}_R, \bar{\varepsilon}_R)$ for each (sufficiently large) $\beta > 0$. We will show that $\varepsilon = \frac{1}{2}(\underline{\varepsilon}_R + \bar{\varepsilon}_R)$.

Using the fact that $f_\beta(E) = 1 - f_\beta(-E)$ for all E , we obtain: for all $\tau \in (\underline{\varepsilon}_R, \bar{\varepsilon}_R)$

$$\begin{aligned} N_{\#}^{\beta,R}(\tau) &= \frac{2}{\#\sigma_R} \left(\sum_{j=1}^{\frac{\#\sigma_R}{2}} \{1 - f_\beta(|\varepsilon_j^R - \tau|)\} + \sum_{j=\frac{\#\sigma_R}{2}+1}^{\#\sigma_R} f_\beta(|\varepsilon_j^R - \tau|) \right) \\ &= 1 + \frac{2}{\#\sigma_R} \left(- \sum_{j=1}^{\frac{\#\sigma_R}{2}} f_\beta(|\varepsilon_j^R - \tau|) + \sum_{j=\frac{\#\sigma_R}{2}+1}^{\#\sigma_R} f_\beta(|\varepsilon_j^R - \tau|) \right). \end{aligned}$$

We therefore obtain the equation

$$\sum_{j=1}^{\frac{\#\sigma_R}{2}} f_\beta(|\varepsilon_j^R - \varepsilon_F^{\beta,R}|) = \sum_{j=\frac{\#\sigma_R}{2}+1}^{\#\sigma_R} f_\beta(|\varepsilon_j^R - \varepsilon_F^{\beta,R}|) \quad (4.7)$$

for each β sufficiently large. Now suppose again that $\varepsilon_F^{\beta,R} \rightarrow \varepsilon \in [\underline{\varepsilon}_R, \bar{\varepsilon}_R]$ along a subsequence as $\beta \rightarrow \infty$. After fixing $\delta > 0$, we know that $\varepsilon_F^{\beta,R}$ is necessarily contained in the interval $(\varepsilon - \delta, \varepsilon + \delta)$ for sufficiently large β . Using (4.7) we obtain the inequalities

$$\underline{m}f_\beta(|\underline{\varepsilon}_R - \varepsilon_F^{\beta,R}|) \leq \frac{\#\sigma_R}{2} f_\beta(|\bar{\varepsilon}_R - \varepsilon_F^{\beta,R}|) \quad \text{and} \quad \bar{m}f_\beta(|\bar{\varepsilon}_R - \varepsilon_F^{\beta,R}|) \leq \frac{\#\sigma_R}{2} f_\beta(|\underline{\varepsilon}_R - \varepsilon_F^{\beta,R}|)$$

which hold for sufficiently large $\beta > 0$. These give rise to the following equations

$$\frac{f_\beta(\varepsilon + \delta - \underline{\varepsilon}_R)}{f_\beta(\bar{\varepsilon}_R - (\varepsilon + \delta))} \leq C \quad \text{and} \quad \frac{f_\beta(\bar{\varepsilon}_R - (\varepsilon - \delta))}{f_\beta((\varepsilon - \delta) - \underline{\varepsilon}_R)} \leq C \quad (4.8)$$

for all β sufficiently large and some β independent constant $C > 0$. Now, the conditions in (4.8) ensure that $\varepsilon + \delta - \underline{\varepsilon}_R \geq \bar{\varepsilon}_R - (\varepsilon + \delta)$ and $\bar{\varepsilon}_R - (\varepsilon - \delta) \geq (\varepsilon - \delta) - \underline{\varepsilon}_R$. These inequalities imply

$$\frac{\underline{\varepsilon}_R + \bar{\varepsilon}_R}{2} - \delta \leq \varepsilon \leq \frac{\underline{\varepsilon}_R + \bar{\varepsilon}_R}{2} + \delta$$

and so, because the choice of $\delta > 0$ was arbitrary, we conclude that $\varepsilon = \frac{1}{2}(\underline{\varepsilon}_R + \bar{\varepsilon}_R)$. $\square$

Step 4. Using the assumptions from Theorem 1, we have

$$\lim_{R \rightarrow \infty} \lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta, R} = \varepsilon_F^{\infty, \#}.$$

$$\begin{array}{ccc} \varepsilon_F^{\beta, R} & \xrightarrow{(1)} & \varepsilon_F^{\beta, \#} \\ (3) \downarrow & \text{~~~~~} (5) \text{~~~~~} & \downarrow (2) \\ \frac{1}{2}(\underline{\varepsilon}_R + \bar{\varepsilon}_R) & \xrightarrow{(4)} & \frac{1}{2}(\underline{\varepsilon} + \bar{\varepsilon}) \end{array}$$

Proof. We shall use the notation $\varepsilon_F^{\infty, R} := \frac{1}{2}(\underline{\varepsilon}_R + \bar{\varepsilon}_R)$. We know that, along a subsequence (which we do not relabel), we can find ε, ℓ such that $\varepsilon_F^{\infty, R} \rightarrow \varepsilon$, $N_{\#}^{\infty, R}(\varepsilon_F^{\infty, R}) \rightarrow \ell$ as $R \rightarrow \infty$. This means that for all $\delta > 0$ we can find a $R' > 0$ such that $\varepsilon - \delta \leq \varepsilon_F^{\infty, R} \leq \varepsilon + \delta$ for all $R > R'$. Since $N_{\#}^{\infty, R}(\cdot)$ is increasing and $N_{\#}^{\infty, R}(\cdot) \rightarrow N_{\#}^{\infty, \#}(\cdot)$ pointwise as $R \rightarrow \infty$, we obtain

$$N_{\#}^{\infty, R}(\varepsilon - \delta) \leq N_{\#}^{\infty, R}(\varepsilon_F^{\infty, R}) \leq N_{\#}^{\infty, R}(\varepsilon + \delta)$$

and thus

$$N_{\#}^{\infty, \#}(\varepsilon - \delta) \leq \ell \leq N_{\#}^{\infty, \#}(\varepsilon + \delta).$$

By continuity of $N_{\#}^{\infty, \#}(\cdot)$ we have that $\ell = N_{\#}^{\infty, \#}(\varepsilon)$. We know

$$N_{\#}^{\infty, R}(\varepsilon_F^{\infty, R}) \in \left(1 - \frac{m}{\#\sigma_R}, 1 + \frac{\bar{m}}{\#\sigma_R}\right)$$

and so $N_{\#}^{\infty, \#}(\varepsilon) = 1$. Now, in the case that we do not have a band gap, $N_{\#}^{\infty, \#}(\cdot)$ is strictly increasing and so ε is the required Fermi level.

Next, we consider the possibility of a band gap ($\underline{\varepsilon} \neq \bar{\varepsilon}$). We remark that, by Proposition 1, we know that $\sigma(\mathcal{H}^R) \subset \text{supp} \rho$ and moreover, $\underline{\varepsilon}_R \leq \underline{\varepsilon}$ and $\bar{\varepsilon}_R \geq \bar{\varepsilon}$ for each $R > 0$. Suppose that $\underline{\varepsilon}_R \rightarrow \varepsilon$ along a subsequence as $R \rightarrow \infty$. If $\varepsilon < \underline{\varepsilon}$, then, for some small $\delta > 0$, we will have $\underline{\varepsilon}_R < \varepsilon + \delta < \underline{\varepsilon}$ for all sufficiently large $R > 0$. This however implies that, for sufficiently large R , there are no eigenvalues of $\mathcal{H}^R$ in the range $(\varepsilon + \delta, \underline{\varepsilon})$ and so $\rho((\varepsilon + \delta, \underline{\varepsilon})) = 0$ contradicting the choice of $\underline{\varepsilon}$. Similarly, $\bar{\varepsilon}_R \rightarrow \bar{\varepsilon}$ as $R \rightarrow \infty$. We have thus shown that

$$\frac{1}{2}(\underline{\varepsilon}_R + \bar{\varepsilon}_R) \rightarrow \frac{1}{2}(\underline{\varepsilon} + \bar{\varepsilon})$$

as $R \rightarrow \infty$. □

4.2 Proof of Theorem 2

We start by proving the interlacing property:

Proof of Proposition 3. We first consider the case of a rank one perturbation. Let A be a real symmetric matrix and B a real symmetric rank one matrix. We will first show that the eigenvalues of $A + B$ interlace those of A :

Lemma 2 (Stated in [3]). *Suppose we have $n \times n$ real symmetric matrices A and B and further suppose that B is of rank one. Then, either*

$$\begin{aligned} \lambda_1(A + B) \leq \lambda_1(A) \leq \dots \leq \lambda_k(A + B) \leq \lambda_k(A) \leq \dots \leq \lambda_n(A + B) \leq \lambda_n(A) \quad \text{or} \\ \lambda_1(A) \leq \lambda_1(A + B) \leq \dots \leq \lambda_k(A) \leq \lambda_k(A + B) \leq \dots \leq \lambda_n(A) \leq \lambda_n(A + B). \end{aligned}$$

For now, we assume Lemma 2. If we instead consider a rank k matrix, B , it is easy to see that B can be written as the sum of a finite number, k' , of rank one matrices. After applying Lemma 2 a finite number of times, we can conclude. $\square$

Proof of Lemma 2. First, assuming $B \neq 0$, we can find a normalised real vector u and some scalar α such that $B = \alpha uu^T$. We simply compute the characteristic polynomial of $A + B$ at a point $\lambda \notin \sigma(A)$:

$$\begin{aligned} p_{A+B}(\lambda) &:= \det(A + \alpha uu^T - \lambda I) \\ &= \det(A - \lambda I) \det(I + \alpha(A - \lambda I)^{-1} uu^T). \end{aligned}$$

Applying the Matrix Determinant Lemma [8], we obtain

$$p_{A+B}(\lambda) = p_A(\lambda) (1 + \alpha u^T (A - \lambda I)^{-1} u).$$

That is, if A has eigenvalues $\lambda_1 \leq \dots \leq \lambda_n$ with associated eigenvectors $(v^{(j)})_j$, then

$$p_{A+B}(\lambda) = p_A(\lambda) \left(1 + \alpha \sum_{j=1}^n \frac{(u \cdot v^{(j)})^2}{\lambda_j - \lambda} \right).$$

Therefore, eigenvalues of $A + B$ are either eigenvalues of A or roots of the equation

$$g(\lambda) := \alpha \sum_{j=1}^n \frac{(u \cdot v^{(j)})^2}{\lambda - \lambda_j} = 1. \tag{4.9}$$

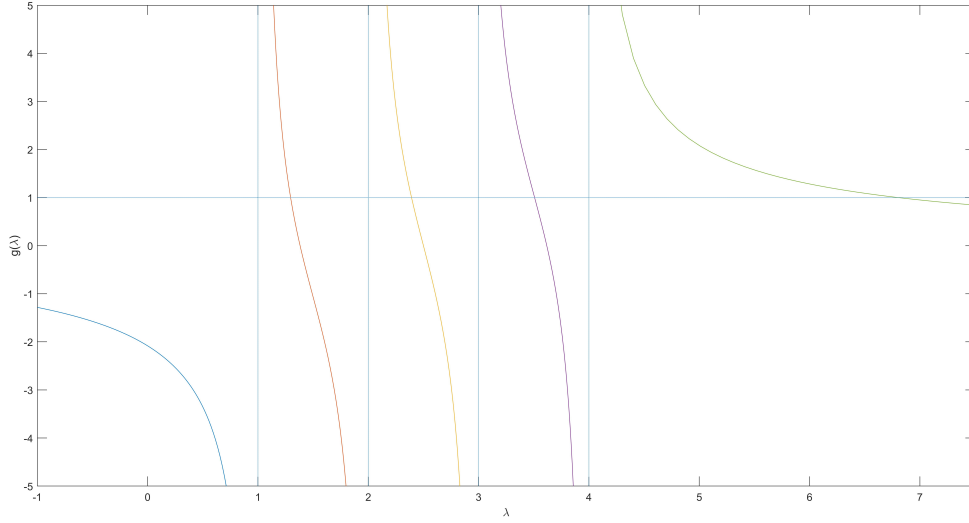


Figure 4.1: Plot of the function $g(\lambda) = \alpha \sum_{j=1}^n \frac{(u \cdot v^{(j)})^2}{\lambda - \lambda_j}$ where A is the 4×4 diagonal matrix with diagonal $(1 \ 2 \ 3 \ 4)$ and $B = \alpha u u^T$ with $\alpha = 1$ and $u = (1 \ 1 \ 1 \ 1)^T$. We can see that solutions to the equation $g(\cdot) = 1$ satisfy the interlacing property described in Lemma 2.

We take $1 \leq j \leq n$ such that λ_j is not an eigenvalue of $A + B$. Since $(A + B)v^{(j)} = \lambda_j v^{(j)} + \alpha (u \cdot v^{(j)}) u$, we can see that $u \cdot v^{(j)} \neq 0$. We define $k \in \mathbb{N} \cup \{\infty\}$ by

$$k := \inf\{m : \lambda_m > \lambda_j \text{ and } u \cdot v^{(m)} \neq 0\}.$$

Therefore $\lambda_m \in \sigma(A + B)$ for each $\lambda_j < \lambda_m < \lambda_k$ and so

$$g(\lambda) \rightarrow \begin{cases} +\infty & \text{as } \lambda \searrow \lambda_j \\ -\infty & \text{as } \lambda \nearrow \lambda_k \text{ if } k < \infty \\ 0 & \text{as } \lambda \rightarrow \infty \text{ if } k = \infty. \end{cases}$$

Moreover, after defining

$$I := \begin{cases} (\lambda_j, \lambda_k) & \text{if } k < \infty \\ (\lambda_j, \infty) & \text{if } k = \infty, \end{cases}$$

we know that, $g(\cdot)$ is a strictly decreasing continuous function on I and thus there is a unique solution to $g(\cdot) = 1$ on I . We will denote this solution by λ^* . Thus, we can see

that for $j \leq m < k$,

$$\lambda_m(A+B) = \begin{cases} \lambda_{m+1} & \text{if } \lambda^* > \lambda_{m+1} \\ \lambda^* & \text{if } \lambda_m \leq \lambda^* \leq \lambda_{m+1} \\ \lambda_m & \text{if } \lambda^* < \lambda_m \end{cases}$$

and thus we have the required interlacing property. The $\alpha < 0$ case can be treated in the exact same way. $\square$

Throughout the proof of Theorem 2 we shall use the notation $\|\cdot\|_F$ for the Frobenius norm and will write

$$\sigma(\mathcal{H}^R(x+u)) = \{\varepsilon_1^R(u) \leq \dots \leq \varepsilon_{\#\sigma_R}^R(u)\}$$

where $\#\sigma_R := \#\sigma(\mathcal{H}^R(x+u))$ of course depends on R but also on the number of orbitals that we are considering.

Step 1. *Using the notation of Theorem 2, we have that for all $\delta > 0$, there exists a matrix $P_1(u)$ with finite R -independent rank such that*

$$\mathcal{H}(x+u) = \mathcal{H}(x) + P_1(u) + P_2(u)$$

and $\|P_2(u)\|_F \leq \delta$.

Proof. Since the $h^{\alpha\beta}$ are C^1 functions, we can apply Taylor's theorem to conclude that for each $\ell, k \in \Lambda$ and $\alpha, \beta \in \Xi$, there exists $\xi_{\ell k}^{\alpha\beta} = (\ell - k) + \theta(u(\ell) - u(k))$ for some $\theta = \theta(\ell, k, \alpha, \beta) \in [0, 1]$ such that

$$h^{\alpha\beta}(y(\ell) - y(k)) = h^{\alpha\beta}(\ell - k) + \nabla h^{\alpha\beta}(\xi_{\ell k}^{\alpha\beta}) \cdot (u(\ell) - u(k)). \quad (4.10)$$

This means that we can write $\mathcal{H}(x+u) = \mathcal{H}(x) + P(u)$ where the perturbation $P(u)$ is given by

$$P(u)_{\ell k}^{\alpha\beta} = \nabla h^{\alpha\beta}(\xi_{\ell k}^{\alpha\beta}) \cdot (u(\ell) - u(k)).$$

Since $x \in \mathcal{V}_m$ and $x+u \in \mathcal{V}_m$, we know that $x+\theta u \in \mathcal{V}_m$. This, together with **(TB)** and the fact that $\xi_{\ell k}^{\alpha\beta} = (x+\theta u)(\ell) - (x+\theta u)(k)$ gives

$$|\nabla h^{\alpha\beta}(\xi_{\ell k})| \leq h_1 e^{-\gamma_1 |\xi_{\ell k}|} \leq h_1 e^{-\gamma_1 m |\ell - k|}. \quad (4.11)$$

Further, as $u \in \text{Adm}_{\mathbf{m}}$, we have

$$\|P(u)\|_{\text{F}} \leq n_{\Xi} h_1 \|u\|_{\gamma} < \infty$$

where $\gamma := \gamma_1 \mathbf{m} > 0$. We will use this to write $P(u)$ as the sum of a matrix of finite R independent rank and a small perturbation in the Frobenius norm.

Fix an arbitrary $\delta > 0$ and choose $R_{\delta} > 0$ such that

$$\|P(u)\|_{\text{F}}^2 - \sum_{\alpha, \beta \in \Xi} \sum_{\ell, k \in \Lambda_{R_{\delta}}} \left(P(u)_{\ell k}^{\alpha \beta} \right)^2 \leq \delta^2$$

Now, we can write the perturbation $P(u)$ as the sum of two contributions $P_1(u)$ and $P_2(u)$ where

$$P(u)_{\ell k}^{\alpha \beta} = \begin{cases} P_1(u)_{\ell k}^{\alpha \beta} & \text{if both } \ell, k \in \Lambda_{R_{\delta}} \\ P_2(u)_{\ell k}^{\alpha \beta} & \text{otherwise.} \end{cases}$$

By the choice of R_{δ} , we know that $\|P_2(u)\|_{\text{F}} \leq \delta$. On the other hand, the rank of $P_1(u)$ is finite and only depends on R_{δ} and the number of orbitals that we are considering. $\square$

We now see how this affects the spectrum of the Hamiltonian. For fixed R , we are considering the spectrum of $\mathcal{H}^R(x+u) = \mathcal{H}^R(x) + P_1(u) + P_2(u)$. Since $P_1(u)$ has some finite rank k' , independent of R , we may apply Proposition 3, to conclude that the spectrum of $\mathcal{H}^R(x) + P_1(u)$ has the interlacing property. That is, after denoting the j^{th} smallest eigenvalues of $\mathcal{H}^R$ and $\mathcal{H}^R + P_1(u)$ by ε_j and $\tilde{\varepsilon}_j$ respectively, we have that $\varepsilon_{j-k'} \leq \tilde{\varepsilon}_j \leq \varepsilon_{j+k'}$ for each $k' < j \leq \#\sigma_R - k'$. Now, by Proposition 4, adding on $P_2(u)$ to the resulting Hamiltonian can only perturb each eigenvalue by at most $\|P_2(u)\|_{\text{F}} \leq \delta$.

Now we see what this description of the perturbed spectrum means for the associated Fermi levels.

Step 2. *Using the assumptions from Theorem 2, for all $\beta > 0$, we have*

$$\lim_{R \rightarrow \infty} \varepsilon_{\text{F}}^{\beta, R}(u) = \varepsilon_{\text{F}}^{\beta, \#}.$$

$$\begin{array}{ccccc} & \varepsilon_{\text{F}}^{\beta, R}(u) & \xrightarrow{\quad} & \varepsilon_{\text{F}}^{\beta, \#} & \xrightarrow{\quad} & \varepsilon_{\text{F}}^{\beta, \#} \\ & \downarrow & & \downarrow & \nearrow & \\ \varepsilon_{\text{F}}^{\beta, R} & \xrightarrow{\quad} & \varepsilon_{\text{F}}^{\beta, \#} & & & \\ \downarrow & & \downarrow & & & \\ \lim_{\beta \rightarrow \infty} \varepsilon_{\text{F}}^{\beta, R}(u) & \xrightarrow{\quad} & \lim_{\beta \rightarrow \infty} \varepsilon_{\text{F}}^{\beta, \#} & \xrightarrow{\quad} & \lim_{R \rightarrow \infty} \lim_{\beta \rightarrow \infty} \varepsilon_{\text{F}}^{\beta, R}(u) & \\ \downarrow & & \downarrow & & & \\ \lim_{R \rightarrow \infty} \varepsilon_{\text{F}}^{\beta, R} & \xrightarrow{\quad} & \lim_{\beta, R \rightarrow \infty} \varepsilon_{\text{F}}^{\beta, R} & \xrightarrow{\quad} & \lim_{\beta, R \rightarrow \infty} \varepsilon_{\text{F}}^{\beta, R} & \\ & & \text{metal} & & & \end{array}$$

Proof. First of all, we note that Step 1 of Theorem 1 still holds in this more general case.

That is, $\varepsilon_{\text{F}}^{\beta, R}(u)$ converges to the solution, τ , of the following equation

$$2 \int f_{\beta}(\varepsilon - \tau) d\rho'(\varepsilon) = 1$$

where ρ' is the density of states corresponding to the perturbed system:

$$\rho'(A) = \lim_{R \rightarrow \infty} \frac{1}{\#\sigma_R} \#(\sigma(\mathcal{H}^R(x+u)) \cap A)$$

for all Borel $A \subset \mathbb{R}$. Therefore, to complete this step, we only need to show that the density of states corresponding to the perturbed Hamiltonian is the same as in the homogeneous case. We now fix some interval $(a, b) \subset \mathbb{R}$ and show that $\rho((a, b)) = \rho'((a, b))$. From the interlacing property (Proposition 3), adding on the finite rank perturbation to the homogeneous spectrum may introduce at most k' eigenvalues into the interval from each end point. On the other hand, by the same argument, no more than $2k'$ eigenvalues can leave the interval. Since k' is independent of R , the density of states corresponding to only adding this finite rank perturbation is unchanged. Now, by introducing the small perturbation matrix, we perturb the spectrum by at most δ . We therefore obtain,

$$\rho((a + \delta, b - \delta)) \leq \rho'((a, b)) \leq \rho((a - \delta, b + \delta)).$$

As $\delta > 0$ was chosen arbitrarily, we can conclude that $\rho = \rho'$. We have thus shown that

$$\lim_{R \rightarrow \infty} \varepsilon_F^{\beta, R}(u) = \varepsilon_F^{\beta, \#}. \quad \square$$

Step 3. *Using the assumptions from Theorem 2, we have that*

$$\lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta, R}(u) \text{ exists.}$$

$$\begin{array}{ccccc} & & \varepsilon_F^{\beta, R}(u) & \xrightarrow{\quad} & \varepsilon_F^{\beta, \#} \\ & & \downarrow & & \nearrow \\ \varepsilon_F^{\beta, R} & \xrightarrow{\quad} & \varepsilon_F^{\beta, \#} & & \\ \downarrow & & \downarrow & & \\ \lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta, R} & \xrightarrow{\quad} & \lim_{\beta, R \rightarrow \infty} \varepsilon_F^{\beta, R} & \xrightarrow{\quad} & \lim_{R \rightarrow \infty} \lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta, R}(u) \\ & & \text{metal} & & \end{array}$$

Proof. We showed in Step 3 of the proof of Theorem 1 that the zero temperature limit of the Fermi level corresponding to the subsystem is the midpoint of the two middle eigenvalues. This is purely based upon the fact we have some finite spectrum and so we can equally apply this result to $\sigma(\mathcal{H}^R(x+u))$. Therefore, we have

$$\varepsilon_F^{\beta, R}(u) \rightarrow \frac{1}{2} (\underline{\varepsilon}_R(u) + \bar{\varepsilon}_R(u)) \quad \text{as } \beta \rightarrow \infty$$

where $\underline{\varepsilon}_R(u) := \varepsilon_{\frac{\#\sigma_R}{2}}(u)$ and $\bar{\varepsilon}_R(u) := \varepsilon_{\frac{\#\sigma_R}{2}+1}(u)$. Therefore $\lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta, R}(u)$ exists but need not be equal to $\frac{1}{2}(\underline{\varepsilon}_R + \bar{\varepsilon}_R)$ due to the introduction of new eigenvalues inside the band gap. $\square$

Step 4. Under the assumptions of Theorem 2, in the case of a metal, the zero temperature and thermodynamic limits of the Fermi level exist and commute:

$$\lim_{R \rightarrow \infty} \lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta,R}(u) = \lim_{\beta \rightarrow \infty} \lim_{R \rightarrow \infty} \varepsilon_F^{\beta,R}(u) = \lim_{\beta, R \rightarrow \infty} \varepsilon_F^{\beta,R}.$$

Proof. We know that $\lim_{\beta \rightarrow \infty} \lim_{R \rightarrow \infty} \varepsilon_F^{\beta,R}(u) = \lim_{\beta, R \rightarrow \infty} \varepsilon_F^{\beta,R}$ by Step 2 and Theorem 1. Further, $\lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta,R}(u)$ exists by Step 3.

Since $\lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta,R}(u) = \frac{1}{2}(\underline{\varepsilon}_R(u) + \bar{\varepsilon}_R(u))$, we only need to show that, in the case of a metal, $\underline{\varepsilon}_R(u), \bar{\varepsilon}_R(u) \rightarrow \underline{\varepsilon} = \bar{\varepsilon}$ as $R \rightarrow \infty$. However, this follows from the description of the perturbed spectrum. Indeed, for all $\delta > 0$, there exists a $k' > 0$ such that

$$\varepsilon_{\frac{\#\sigma_R}{2}-k'}^R - \delta \leq \underline{\varepsilon}_R(u) \leq \varepsilon_{\frac{\#\sigma_R}{2}+k'}^R + \delta \quad \text{and} \quad \varepsilon_{\frac{\#\sigma_R}{2}+1-k'}^R - \delta \leq \bar{\varepsilon}_R(u) \leq \varepsilon_{\frac{\#\sigma_R}{2}+1+k'}^R + \delta.$$

Taking $R \rightarrow \infty$, we obtain

$$\frac{1}{2}(\underline{\varepsilon} + \bar{\varepsilon}) - \delta \leq \lim_{R \rightarrow \infty} \lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta,R}(u) \leq \frac{1}{2}(\underline{\varepsilon} + \bar{\varepsilon}) + \delta.$$

Since $\delta > 0$ was chosen arbitrarily, we can conclude. $\square$

Step 5. Under the assumptions of Theorem 2, the limit $\lim_{R \rightarrow \infty} \lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta,R}(u)$ exists.

$$\begin{array}{ccccc} & & \varepsilon_F^{\beta,R}(u) & \xrightarrow{\quad} & \varepsilon_F^{\beta,\#} \\ & & \downarrow & & \searrow \\ \varepsilon_F^{\beta,R} & \xrightarrow{\quad} & \varepsilon_F^{\beta,\#} & & \\ \downarrow & & \downarrow & & \downarrow \\ \lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta,R} & \xrightarrow{\quad} & \lim_{\beta, R \rightarrow \infty} \varepsilon_F^{\beta,R} & \xrightarrow{\quad} & \lim_{\beta, R \rightarrow \infty} \varepsilon_F^{\beta,R} \text{ (metal)} \\ & & \uparrow & & \uparrow \\ & & \lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta,R}(u) & \xrightarrow{\quad} & \lim_{R \rightarrow \infty} \lim_{\beta \rightarrow \infty} \varepsilon_F^{\beta,R}(u) \end{array}$$

Proof. Recall, we only prove this step for isolated eigenvalues in the band gap. For the case of a metal, this has been shown in Step 4 and so we assume that there is a band gap in the spectrum.

We will simply show that the newly introduced eigenvalues in the band gap converge as $R \rightarrow \infty$. We will use the following lemma:

Lemma 3 (Lemma 2.2 [6]). Suppose that $\mathcal{H}^R$ is a Hamiltonian satisfying **TB** and let $\mathcal{C} \subset \mathbb{C}$ be a curve in the complex plane such that

$$\text{dist}(\mathcal{C}, \sigma(\mathcal{H}^R)) \geq d$$

for some $d > 0$ independent of R . Then, there exists a positive constant $\gamma_2 > 0$, independent of R , such that

$$|(\mathcal{H}^R - z)_{\ell k}^{-1}| \leq 2d^{-1}e^{-\gamma_2|\ell-k|}$$

for all $z \in \mathcal{C}$ and each R .

We suppose that there exists a single eigenvalue in the interior of the band gap for each R and denote this eigenvalue by ε_R . The band gap is a compact interval and so we know there exists a convergent subsequence. We suppose there are two such subsequences $(R_n)_n, (R'_n)_n$ such that $\varepsilon_{R_n} \rightarrow \varepsilon$ and $\varepsilon_{R'_n} \rightarrow \varepsilon'$ as $n \rightarrow \infty$ where $\varepsilon \neq \varepsilon'$ and $\varepsilon, \varepsilon'$ are contained in the interior of the band gap. This means we can find a (R -independent) curve $\mathcal{C} \subset \mathbb{C}$ encircling the eigenvalues ε_R and a positive constant $\delta > 0$ such that

$$\delta < \text{dist}(\sigma(\mathcal{H}^R), \mathcal{C}) \quad \text{and} \quad \delta \leq \frac{1}{4}|\varepsilon - \varepsilon'|.$$

By removing finitely many terms from each sequence, we further suppose that $\varepsilon_{R_n} \in B(\varepsilon; \delta)$ and $\varepsilon_{R'_n} \in B(\varepsilon'; \delta)$ for all n .

We will show that the eigenvector corresponding to ε_R (for $R \in \{R_n\}_n$) decays exponentially and thus will show that it is “close” to an eigenvector corresponding to ε_S on the larger domain $\Lambda_S \supset \Lambda_R$ such that $S \in \{R'_n\}_n$. We will then use the Inverse Function Theorem to conclude that there exists a nearby eigenvector with eigenvalue ε_R , a contradiction.

We let ψ_R be an eigenvector of $\mathcal{H}^R$ corresponding to ε_R and use the eigen-decomposition of $\mathcal{H}^R$ to show that the eigenvector has exponential decay:

$$\frac{1}{2\pi i} \oint_{\mathcal{C}} (z - \mathcal{H}^R)^{-1} dz = \psi_R \otimes \psi_R. \quad (4.12)$$

Now, by applying Lemma 3 we have: for each $\ell, k \in \Lambda_R$

$$|\psi_R(\ell)\psi_R(k)| \leq C_1 e^{-\gamma_2|\ell-k|}$$

for some positive (R -independent) constant C_1 . We therefore obtain the required exponential decay of the eigenvector:

$$|\psi_R(\ell; \alpha)| \leq C_1 e^{-\gamma_2|\ell|} \quad (4.13)$$

for all $\alpha \in \Xi$ and $\ell \in \Lambda_R$.

For given R and $R_{\text{buf}} < R$ let us define the sets

$$\Lambda_R^\circ := \Lambda_{R-R_{\text{buf}}} \quad \text{and} \quad \partial\Lambda_R := \Lambda_R \setminus \Lambda_R^\circ.$$

Now for a given $S > R$ we define $\tilde{\psi}_R$ on Λ_S by

$$\tilde{\psi}_R(\ell; \alpha) := \begin{cases} \psi_R(\ell; \alpha) & \text{if } \ell \in \Lambda_R^\circ \\ 0 & \text{otherwise.} \end{cases}$$

If $\ell \in \Lambda_S \setminus \Lambda_R$ then, by **(TB)**, we have

$$\begin{aligned} \left| \mathcal{H}^S \tilde{\psi}_R(\ell; \alpha) \right| &= \left| \sum_{\beta \in \Xi} \sum_{k \in \Lambda_R^\circ} (\mathcal{H}^S)_{\ell k}^{\alpha\beta} \psi(k; \beta) \right| \\ &\leq C e^{-\gamma_1 m R_{\text{buf}}/2} \end{aligned}$$

for some $C > 0$ independent of R, R_{buf} and S . Moreover, for $\ell \in \Lambda_R$, we have

$$\begin{aligned} \mathcal{H}^S \tilde{\psi}(\ell; \alpha) &= \sum_{\beta \in \Xi} \sum_{k \in \Lambda_R^\circ} (\mathcal{H}^S)_{\ell k}^{\alpha\beta} \psi(k; \beta) \\ &= \varepsilon_R \psi(\ell; \alpha) - \sum_{\beta \in \Xi} \sum_{k \in \partial \Lambda_R} (\mathcal{H}^S)_{\ell k}^{\alpha\beta} \psi(k; \beta). \end{aligned}$$

We can therefore write

$$\mathcal{H}^S \tilde{\psi}_R(\ell; \alpha) = \varepsilon_R \tilde{\psi}(\ell; \alpha) + \eta_{R, R_{\text{buf}}}(\ell; \alpha)$$

where

$$|\eta_{R, R_{\text{buf}}}(\ell; \alpha)| \leq \tilde{C} e^{-\gamma \max\{R_{\text{buf}}, R - R_{\text{buf}}\}}$$

for some positive constants $\tilde{C}, \gamma > 0$ independent of R, R_{buf} and S .

At this point we introduce the Inverse Function Theorem [17, Theorem 2.1]:

Inverse Function Theorem. *Suppose $F: X \rightarrow Y$ is Fréchet differentiable at a point $x_0 \in X$ where X, Y are two Banach spaces. Further, suppose that $DF(x_0)$ is invertible and that there exists strictly positive constants $\eta, \sigma, L > 0$ such that the following holds*

$$\begin{aligned} 2L\sigma^2\eta &< 1 \\ \|F(x_0)\|_Y &\leq \eta \\ \|DF(x_0)^{-1}\|_{op} &\leq \sigma \\ \|DF(x) - DF(x')\|_{op} &\leq L\|x - x'\|_X \quad \text{for all } x, x' \in B(x_0, 2\sigma\eta). \end{aligned}$$

Then there exists a unique $x \in B(x_0, 2\sigma\eta)$ such that $F(x) = 0$.

We now fix $R = R_n$ and R_{buf} sufficiently large so that

$$\tilde{C} e^{-\gamma \max\{R_{\text{buf}}, R - R_{\text{buf}}\}} \leq \frac{1}{4} \delta^2$$

and $S = R'_m > R$ for some m . To obtain a contradiction, we will now show that ε_R is an eigenvalue of $\mathcal{H}^S$. We introduce the function F defined by

$$F(\phi) := \mathcal{H}^S \phi - \varepsilon_R \phi.$$

Notice that zeros ϕ^* of F such that $\|\phi^*\| \neq 0$ correspond to eigenvectors with eigenvalue ε_R . By our choice of R and R_{buf} , we know $\|F(\tilde{\psi}_R)\|_\infty \leq \frac{1}{4}\delta^2$. Moreover, differentiating, we have

$$DF(\phi) = \mathcal{H}^S - \varepsilon_R I$$

which clearly satisfies the Lipschitz condition from the Inverse Function Theorem. By choice of δ , we know that

$$\|DF(\psi_R^S)^{-1}\| \leq \delta^{-1}.$$

Now, the condition $2L\sigma^2\eta < 1$ is satisfied with $L = 1, \sigma = \delta^{-1}, \eta = \frac{1}{4}\delta^2$ and so, there is a zero ϕ^* of F satisfying $\|\phi^* - \tilde{\psi}_R\| \leq \frac{1}{2}$. In particular, as $\tilde{\psi}_R$ is normalised, we have that ϕ^* is non-zero and thus is an eigenvector of $\mathcal{H}^S$ with corresponding eigenvalue ε_R . This implies $\varepsilon_S = \varepsilon_R$, a contradiction. $\square$

Remark 5. *We again stress that we do not prove Theorem 2 in full generality. We have only shown that Step 5 holds in the case of single isolated eigenvalues in the band gap (as was the case in Example 4). This can be extended to multiple isolated eigenvalues in the band gap. In this case, the projector in (4.12) will now take the form $\sum_j \psi_R^{(j)} \otimes \psi_R^{(j)}$ and similar arguments can be made. If the eigenvalues cluster as $R \rightarrow \infty$, or we have repeated eigenvalues, more arguments are needed.*

Bibliography

- [1] Rajendra Bhatia. Linear algebra to quantum cohomology: The story of alfred horn's inequalities. *The American Mathematical Monthly*, 108(4):289–318, 2001.
- [2] P. J. Bickel and P. W. Millar. Uniform convergence of probability measures on classes of functions. *Statistica Sinica*, 2(1):1–15, 1992.
- [3] James R. Bunch, Christopher P. Nielsen, and Danny C. Sorensen. Rank-one modification of the symmetric eigenproblem. *Numerische Mathematik*, 31(1):31–48, 1978.
- [4] H Chen, M Luskin, C Ortner, and A Shapeev. *Multiscale Materials Simulation at the Atomic Scale: Numerical Analysis of Mechanical Embedding Methods*. Manuscript in preparation.
- [5] Huajie Chen, Jianfeng Lu, and Christoph Ortner. Thermodynamic limit of crystal defects with finite temperature tight binding, 2016.
- [6] Huajie Chen and Christoph Ortner. QM/MM methods for crystalline defects. Part 1: Locality of the tight binding model, 2015.
- [7] Huajie Chen and Christoph Ortner. Qm/mm methods for crystalline defects. part 2: Consistent energy and force-mixing. 15, 09 2015.
- [8] Jiu Ding and Aihui Zhou. Eigenvalues of rank-one updated matrices with some applications. *Applied Mathematics Letters*, 20(12):1223 – 1226, 2007.
- [9] Mike Finnis. *Interatomic Forces in Condensed Matter (Oxford Series on Materials Modelling)*. Oxford University Press, 2004.

- [10] Gary Harris and Clyde Martin. Shorter notes: The roots of a polynomial vary continuously as a function of the coefficients. *Proceedings of the American Mathematical Society*, 100(2):390–392, 1987.
- [11] Walter A. Harrison. *Electronic structure and the properties of solids: the physics of the chemical bond*. Dover Books on Physics. Dover Publications, 1989.
- [12] Jurgen Jost. *Postmodern analysis*. Springer, 2005.
- [13] Efthimios Kaxiras. *Atomic and Electronic Structure of Solids*. Cambridge University Press, 2003.
- [14] Charles Kittel. *Introduction to Solid State Physics*. Wiley, 2004.
- [15] P.O. Löwdin. On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. *The Journal of Chemical Physics*, 18:365, 1950.
- [16] Richard M. Martin. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, 2004.
- [17] Christoph Ortner. A posteriori existence in numerical computations, 2008.
- [18] Walter Rudin. *Principles of Mathematical Analysis*. International series in pure and applied mathematics. McGraw-Hill, 1976.
- [19] J. C. Slater and G. F. Koster. Simplified LCAO method for the periodic potential problem. *Phys. Rev.*, 94:1498–1524, Jun 1954.
- [20] Adrian P. Sutton. *Electronic Structure of Materials*. Oxford University Press, 1993.