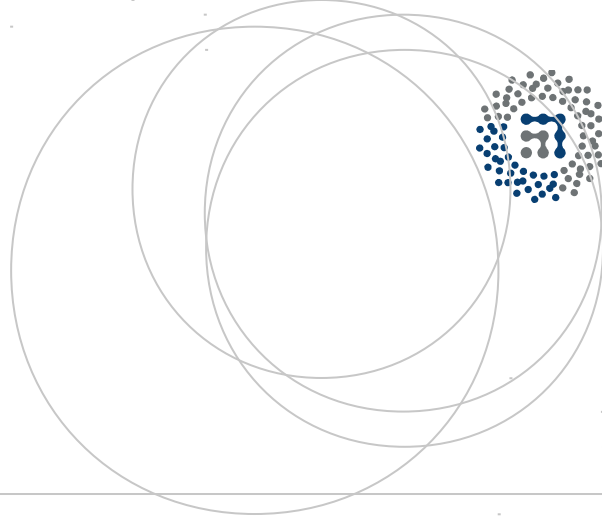


eman ta zabal zazu



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea



ZTF-FCT

Zientzia eta Teknologia Fakultatea
Facultad de Ciencia y Tecnología



Gradu Amaierako Lana / Trabajo Fin de Grado
Fisikako Gradua / Grado en Física

Discusión sobre alternativas al cálculo de funciones de Wannier óptimamente localizadas

Egilea/Autor/a:
Mikel Rouco Martín
Zuzendaria/Director/a:
Asier Eiguren Goyenechea
Idoia García de Gurtubay

Índice

1. Introducción y objetivos	3
2. Fundamento teórico	4
2.1. El problema del sólido	4
2.1.1. Electrón independiente	4
2.2. Electrones en un potencial periódico	4
2.2.1. Operador de traslación, $T_{\mathbf{R}}$	4
2.2.2. Teorema de Bloch	4
2.3. Teoría del funcional de la densidad (DFT)	5
2.4. Estructura de bandas	8
3. Funciones de Wannier óptimamente localizadas (FWOL): Desarrollo teórico	9
3.1. Operador posición proyectado, $\hat{\mathbf{r}}_{\alpha}$	9
3.2. Cumplimiento de las propiedades de las funciones de Wannier	10
3.2.1. Base completa de estados	10
3.2.2. Traslaciones de red	10
3.2.3. Máxima localización	12
3.3. Búsqueda de funciones de Wannier en 3D	13
3.3.1. No conmutabilidad	13
3.3.2. Diagonalización simultánea aproximada. Ángulos de Jacobi.	14
4. Método computacional y resultados	16
4.1. Parámetros del problema	16
4.2. <i>Quantum Espresso</i> para obtener las ondas de Bloch	17
4.2.1. Cálculo autoconsistente	18
4.2.2. Cálculo de bandas	18
4.2.3. Cálculo no-autoconsistente	18
4.3. Referencia: Método de Marzari-Vanderbilt	19
4.4. Obtención de las funciones de Wannier óptimamente localizadas	19
4.4.1. Método de Kivelson	20
4.4.2. Método de diagonalización simultánea	23
5. Conclusiones	29

1. Introducción y objetivos

En 1937, el físico suizo Gregory Wannier introdujo un conjunto de funciones ortogonales en el contexto de la física del estado sólido, ahora conocidas como *funciones de Wannier* [9]. Estas funciones, a diferencia de las ondas de Bloch, se encuentran localizadas en el espacio. Sin embargo, esta localización espacial implica la pérdida de localización en la energía: las funciones de Wannier no son estados propios del hamiltoniano electrónico. Estas propiedades han hecho de las funciones de Wannier una herramienta excelente para el desarrollo de teorías de transporte de los electrones de Bloch o para el estudio de fenómenos que involucren niveles electrónicos localizados (i.e. niveles donadores y aceptores en semiconductores), entre muchos otros. La definición original de estas funciones es la siguiente:

$$\psi_{n\mathbf{R}}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}} \phi_{n\mathbf{k}}(\mathbf{r}) \quad (1)$$

Donde \mathbf{R} es cualquier vector de la red de Bravais, N es el número de celdas unidad que compone el cristal perfecto, $\phi_{\mathbf{k}}(\mathbf{r})$ son ondas de Bloch y la suma se realiza sobre momentos cristalinos pertenecientes a una misma rama electrónica, n . Sin embargo, las ondas de Bloch tienen una indeterminación en una fase, $e^{i\theta_{\mathbf{k}}}$, (simetría *gauge*) que podemos presentar explícitamente en (Ec. 1):

$$\psi_{n\mathbf{R}}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}} \phi_{n\mathbf{k}}(\mathbf{r}) e^{i\theta_{\mathbf{k}}} \quad (2)$$

Dependiendo de la elección de la fase, se pueden construir funciones de Wannier siguiendo distintos criterios. Entre éstos, se encuentran las funciones de Wannier óptimamente (o máximamente) localizadas. En 1997, Marzari y Vanderbilt [6] desarrollaron un método para determinar el conjunto de funciones de Wannier máximamente localizadas asociadas a un conjunto de ondas de Bloch. Este método se basa en un proceso variacional que minimiza el funcional que representa la dispersión total,

$$[\sigma] \equiv \sum_{n\mathbf{R}} \langle n\mathbf{R} | \hat{\mathbf{r}}^2 | n\mathbf{R} \rangle - \langle n\mathbf{R} | \hat{\mathbf{r}} | n\mathbf{R} \rangle^2, \quad (3)$$

donde $\langle \mathbf{r} | n\mathbf{R} \rangle = \psi_{n\mathbf{R}}(\mathbf{r})$. Este método fue implementado en el software *wannier90* [7] con éxito, posibilitando así la obtención de este tipo de funciones mediante cálculos *ab initio*. El método de Marzari-Vanderbilt, sin embargo, conlleva gran dificultad a la hora de implementarlo en un programa, por lo que en este trabajo hemos desarrollado un posible procedimiento alternativo.

En nuestro trabajo, nos hemos basado en el método de Kivelson [5] para el cálculo de funciones de Wannier óptimamente localizadas (FWOL). Kivelson planteó diagonalizar el vector posición proyectado a un subespacio del espacio recíproco para un cristal unidimensional. Nosotros hemos extendido este método al caso tridimensional, donde el operador posición está compuesto por tres operadores, uno por cada componente.

Nuestros objetivos son demostrar que los estados propios de este operador son FWOL y desarrollar un código para resolver el problema de autoestados. Finalmente, los resultados obtenidos mediante nuestro método se contrastarán con los obtenidos mediante el método de Marzari-Vanderbilt.

2. Fundamento teórico

2.1. El problema del sólido

Para estudiar las propiedades electrónicas de un sólido a nivel microscópico es necesario recurrir a la teoría cuántica. En principio, nuestro objetivo sería resolver el sistema de ecuaciones de Schrödinger y obtener los estados propios de los electrones en el material. En nuestro caso no consideramos las vibraciones de red, por lo que el resultado dependería de las posiciones fijas de los átomos en el cristal y de la dinámica del resto de electrones.

2.1.1. Electrón independiente

El problema planteado sería del orden de 10^{23} grados de libertad por centímetro cúbico, pero en la práctica se considera cada electrón de manera independiente; esto es, que no interactúan entre sí. La validez de este método de resolución, que históricamente se tomó como una aproximación, quedó demostrada en 1964 por Hohenberg y Kohn durante el desarrollo de la *teoría del funcional de la densidad*.

2.2. Electrones en un potencial periódico

Al tratar los electrones de manera independiente, nos encontramos ante un hamiltoniano del siguiente tipo:

$$\mathcal{H} = -\frac{\hbar^2}{2m}\nabla^2 + V_{\text{eff}}(\mathbf{r}) \quad (4)$$

donde $V_{\text{eff}}(\mathbf{r})$ es el denominado potencial efectivo que tiene las mismas simetrías que la estructura cristalina del sólido. La forma exacta de este potencial, como veremos en la sección 2.3, no es trivial, ya que no solo contiene la interacción con los iones de la red. La simple simetría bajo traslaciones de red nos da cierta información sobre los estados propios del hamiltoniano.

2.2.1. Operador de traslación, $T_{\mathbf{R}}$

Se define el operador de traslación, $T_{\mathbf{R}}$, como aquel que traslada el argumento de la función sobre la que opera en un vector \mathbf{R} :

$$T_{\mathbf{R}}f(\mathbf{r}) = f(\mathbf{r} + \mathbf{R}) \quad (5)$$

El vector de traslación, \mathbf{R} , puede ser cualquier vector del espacio real. No obstante, en nuestro caso siempre va a ser un vector de red. De ahora en adelante se entiende que $T_{\mathbf{R}}$ traslada el argumento de la función sobre la que actúe en un vector, \mathbf{R} , perteneciente a la red de Bravais.

2.2.2. Teorema de Bloch

Este teorema demuestra que pueden elegirse los autoestados, ψ , de un hamiltoniano electrónico como el de (Ec. 4) (donde $T_{\mathbf{R}}V_{\text{eff}}(\mathbf{r}) = V_{\text{eff}}(\mathbf{r})$) de modo que tengan la forma de una onda plana multiplicada por una función con la periodicidad de la red:

$$\phi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{\mathbf{k}}(\mathbf{r}) \quad (6)$$

donde el vector \mathbf{k} es el llamado *momento cristalino*, perteneciente al *espacio recíproco*. Asimismo, ha de cumplirse que:

$$u_{\mathbf{k}}(\mathbf{r} + \mathbf{R}) = u_{\mathbf{k}}(\mathbf{r}) \quad (7)$$

De manera alternativa, el teorema de Bloch puede enunciarse como los estados propios comunes entre el hamiltoniano, \mathcal{H} , con valores propios $\varepsilon(\mathbf{k})$ y el operador de traslaciones de la red, $T_{\mathbf{R}}$, con valores propios asociados $e^{i\mathbf{k}\cdot\mathbf{R}}$:

$$\boxed{T_{\mathbf{R}}\phi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{R}}\phi_{\mathbf{k}}(\mathbf{r})} \quad (8)$$

para todo \mathbf{R} perteneciente a la red de Bravais.

De esta definición se puede deducir que, siendo \mathbf{G} un vector de la red recíproca:

$$T_{\mathbf{R}}\phi_{\mathbf{k}+\mathbf{G}}(\mathbf{r}) = e^{i\mathbf{G}\cdot\mathbf{R}}e^{i\mathbf{k}\cdot\mathbf{R}}\phi_{\mathbf{k}+\mathbf{G}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{R}}\phi_{\mathbf{k}+\mathbf{G}}(\mathbf{r}) \quad (9)$$

donde se observa que ondas de Bloch asociadas a momentos cristalinos separados por un vector de la red recíproca tienen el mismo valor propio ante el operador de traslación. Por esta razón, se dice que dichos momentos cristalinos son equivalentes y a la hora de representar las energías de los distintos estados en función de \mathbf{k} , se trasladan todos los momentos cristalinos a la denominada *primera zona de Brillouin* (que es el análogo a la celda primitiva de la red de Bravais en la red recíproca).

Para distinguir entre ondas de Bloch separadas por un vector de la red recíproca, añadimos un subíndice, n , a la notación, llamado *índice de rama*. Las ondas de Bloch, de forma más general, se describen de la siguiente manera:

$$\boxed{\phi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{n\mathbf{k}}(\mathbf{r})} \quad (10)$$

2.3. Teoría del funcional de la densidad (DFT)

Para calcular las ondas de Bloch con las que trabajaremos hemos utilizado el software *Quantum Espresso* [3]. Este programa calcula estos niveles electrónicos por medio de la denominada *teoría del funcional de la densidad*, iniciada en 1964 por Hohenberg y Kohn.

En 1964 ambos científicos probaron el siguiente teorema: *Todos los aspectos de la estructura electrónica de un sistema de electrones interactuantes en su estado fundamental, dentro de un potencial 'externo' $v(\mathbf{r})$, se encuentran totalmente determinadas por la densidad de carga electrónica $\rho(\mathbf{r})$* . La enorme simplificación de este resultado radica en que la densidad de carga electrónica es función de solo 3 variables, alejándonos así de las $3N$ variables de posición iniciales. Sin embargo, se trata de una teoría con la que, en principio, tan solo se puede calcular el estado fundamental. Posteriores trabajos han desarrollado técnicas que permiten extender la teoría a cálculos de estados excitados, pero no profundizaremos en ellas.

Nuestro primer objetivo será enunciar la energía del sistema en función de la densidad de carga electrónica, $\rho(\mathbf{r})$, y diversas contribuciones al potencial efectivo. El potencial electrostático debido a la distribución de la densidad de carga electrónica se llama *potencial de Hartree*, $V_H(\mathbf{r})$, y puede obtenerse de la *ecuación de Poisson*:

$$\nabla^2 V_H(\mathbf{r}) = -\frac{\rho(\mathbf{r})}{\epsilon_0} \quad \Longrightarrow \quad V_H(\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' \quad (11)$$

Mientras que el potencial electrostático generado por los iones en el cristal, $V_N(\mathbf{r})$, viene descrito por:

$$V_N(\mathbf{r}) = \sum_i \frac{Z_i e}{4\pi\epsilon_0|\mathbf{r}-\mathbf{R}_i|} \quad (12)$$

donde Z_i es el número de protones del i -ésimo núcleo y \mathbf{R}_i su posición.

Podría pensarse que la energía del sistema vendría dada únicamente por:

$$E[\rho(\mathbf{r})] = T[\rho(\mathbf{r})] + \int \rho(\mathbf{r})V_N(\mathbf{r})d\mathbf{r} + \frac{1}{2} \int \rho(\mathbf{r})V_H(\mathbf{r})d\mathbf{r} \quad (13)$$

donde $T[\rho(\mathbf{r})]$ es el funcional de la energía cinética.

En esta expresión, sin embargo, la interacción Coulombiana se encuentra sobrevalorada por el término de *Hartree* $\frac{1}{2} \int \rho(\mathbf{r})V_H(\mathbf{r})d\mathbf{r}$. Esto se debe a que, por un lado, este término contiene la interacción de un electrón consigo mismo y que, además, no tiene en cuenta el efecto de apantallamiento conocido como *agujero de intercambio-correlación*. Este efecto se entiende como la repulsión entre los electrones por medio de dos interacciones: la de *correlación*, referente a la repulsión electrostática y la de intercambio, en alusión al principio de exclusión de Pauli, que se refiere a la repulsión de electrones cercanos con la misma orientación de espín. En consecuencia, se crea un agujero en la densidad de carga electrónica alrededor de cada electrón, tal y como se representa en la *Figura 1*.

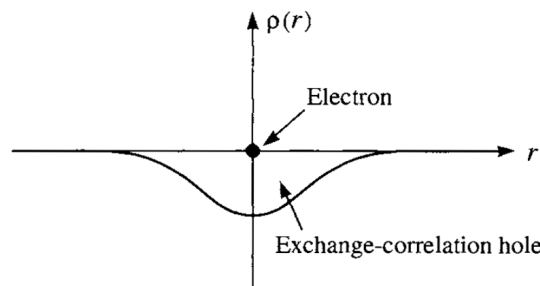


Figura 1: La disminución de la densidad de carga electrónica en los alrededores de cada electrón denominado agujero de intercambio-correlación. Figura tomada de la referencia [8].

En 1965, Kohn y Sham mostraron que, al funcional de la energía, $E[\rho(\mathbf{r})]$, de (Ec. 13), se le ha de sumar el funcional de la *energía de intercambio-correlación*, $E_{XC}[\rho(\mathbf{r})]$. El problema que se plantea es que el funcional, E_{XC} , es desconocido para sistemas en los que la densidad de carga electrónica varía en el espacio (que es lo usual). Sin embargo, para sistemas en los que $\rho(\mathbf{r})$ es uniforme, la energía de *intercambio-correlación* puede

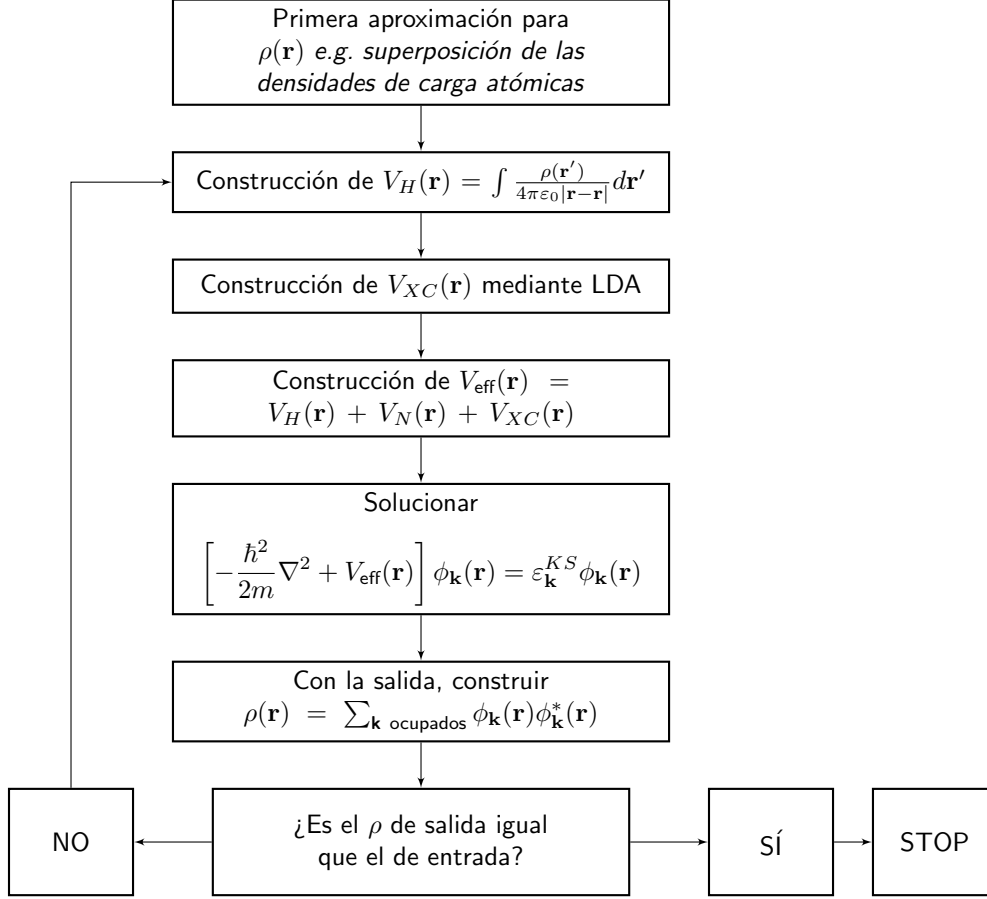


Figura 2: Algoritmo para calcular la densidad de carga electrónica mediante las ecuaciones de Kohn-Sham de manera autoconsistente

calcularse con exactitud. En la práctica, una de las aproximaciones más utilizadas es la llamada LDA (*Local Density Approximation*). En esta aproximación el cristal se discretiza en pequeños elementos de volumen, dentro de los cuales la densidad de carga electrónica se presupone uniforme. Esta es la manera que hemos seleccionado para que opere *Quantum Espresso*, aunque el programa ofrece opciones alternativas.

Por lo tanto, la energía del sistema es funcional de la densidad de carga electrónica, $\rho(\mathbf{r})$, de la siguiente manera:

$$E[\rho(\mathbf{r})] = T[\rho(\mathbf{r})] + \int \rho(\mathbf{r})V_N(\mathbf{r})d\mathbf{r} + \frac{1}{2} \int \rho(\mathbf{r})V_H(\mathbf{r})d\mathbf{r} + E_{XC}[\rho(\mathbf{r})] \quad (14)$$

Siguiendo el proceso típico del método variacional, al minimizar (Ec. 14) junto a la ligadura del número constante de electrones,

$$\int \rho(\mathbf{r})d\mathbf{r} = N, \quad (15)$$

se obtienen las conocidas como *ecuaciones de Kohn-Sham*:

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + V_{\text{eff}}(\mathbf{r}) \right] \phi_{\mathbf{k}}(\mathbf{r}) = \varepsilon_{\mathbf{k}}^{KS} \phi_{\mathbf{k}}(\mathbf{r}) \quad (16)$$

$$V_{\text{eff}}(\mathbf{r}) = V_H(\mathbf{r}) + V_N(\mathbf{r}) + V_{XC}(\mathbf{r}) \quad (17)$$

$$\rho(\mathbf{r}) = \sum_{\mathbf{k} \text{ ocupados}} \phi_{\mathbf{k}}(\mathbf{r})\phi_{\mathbf{k}}^*(\mathbf{r}) \quad (18)$$

El modo de utilizar estas ecuaciones es mediante un algoritmo similar al que se expone en la *Figura 2*. En él se predice una densidad electrónica inicial que se va a recalcular iterativamente utilizando estas ecuaciones hasta su convergencia. De este modo, obtenemos tanto la densidad de carga electrónica, $\rho(\mathbf{r})$, como las ondas de Bloch, $\phi_{\mathbf{k}}(\mathbf{r})$.

2.4. Estructura de bandas

Una vez resuelto el problema electrónico del cristal, podemos representar la energía de cada estado en función del momento cristalino, \mathbf{k} . Como generalmente disponemos de un cristal tridimensional, se definen líneas entre los denominados *puntos de alta simetría* (ciertos \mathbf{k} dentro de la primera zona de Brillouin a los cuales se les asocian letras por convenio) a lo largo de las cuales se representa $\varepsilon(\mathbf{k})$. De este modo, se logra una representación en un gráfico de dos ejes.

Utilizando *Quantum Espresso*, hemos realizado este cálculo para el sistema del silicio con estructura cristalina del tipo diamante (CD), obteniendo así la *Figura 3*.

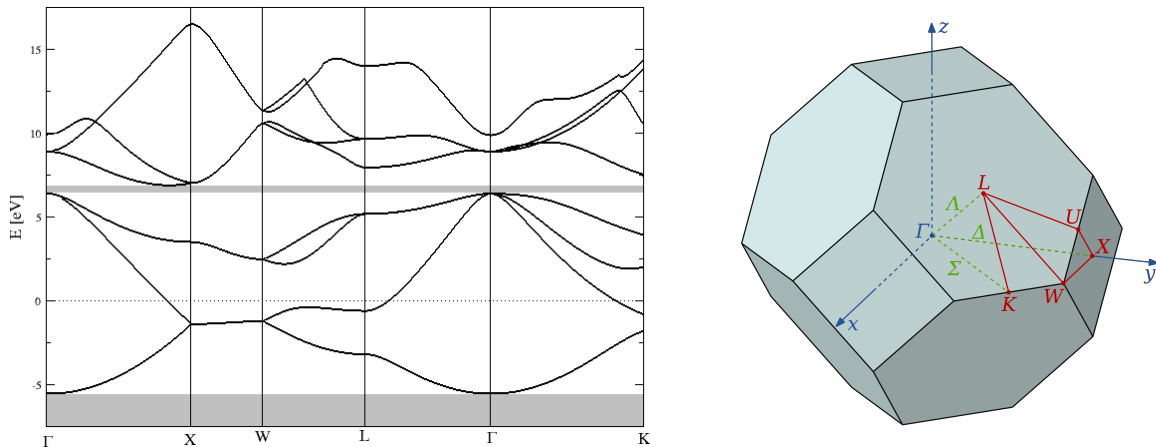


Figura 3: A la izquierda el diagrama de bandas del silicio con estructura de diamante (CD) con las bandas prohibidas sombreadas en gris. A la derecha el convenio utilizado para nombrar los puntos de alta simetría de la primera zona de Brillouin de dicha estructura.

Lo más relevante de la figura es observar que existen secciones de energía en las que hay estados electrónicos (por lo que puede haber electrones con dicha energía) y zonas de energía en las que no se encuentra ningún estado, que se denominan *gaps de energía* (bandas prohibidas grises de la *Figura 3*). A las zonas de posibles energías que se encuentren entre zonas prohibidas se les conoce como *bandas*. Para referirnos a una sola de las líneas continuas de la figura (todas las \mathbf{k} separadas del origen por el mismo vector de la red recíproca) las denominaremos *ramas*.

3. Funciones de Wannier óptimamente localizadas (FWOL): Desarrollo teórico

Debido a la indeterminación en una fase, $e^{i\theta_{\mathbf{k}}}$, de las ondas de Bloch, tenemos libertad en el momento de definir las funciones de Wannier:

$$|n\mathbf{R}\rangle = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}} |n\mathbf{k}\rangle e^{i\theta_{\mathbf{k}}} \quad (19)$$

Uno de los criterios más comunes es el de construir una base de funciones de Wannier máximamente localizadas; esto es, que minimice la dispersión total del conjunto en el espacio real,

$$[\sigma] \equiv \sum_{n\mathbf{R}} \langle n\mathbf{R} | \hat{\mathbf{r}}^2 | n\mathbf{R} \rangle - \langle n\mathbf{R} | \hat{\mathbf{r}} | n\mathbf{R} \rangle^2. \quad (20)$$

Para obtener estas funciones, es común utilizar el software *wannier90* [7]. Este programa utiliza el método de Marzari-Vanderbilt [6] para obtener los valores de $\theta_{\mathbf{k}}$ mediante un proceso variacional que minimiza $[\sigma]$.

En nuestro trabajo, partiendo de un artículo publicado por Kivelson [5] en 1982, desarrollaremos un método alternativo para obtener la FWOL. Las obtendremos a partir de la diagonalización del operador de posición proyectado a un subespacio, α , compuesto por una o más ramas electrónicas, $\hat{\mathbf{r}}_{\alpha}$. Kivelson propuso su método para un subespacio compuesto por una banda de energía aislada en un cristal unidimensional. Nosotros, sin embargo, vamos a extenderlo a cristales tridimensionales y subespacios, α , generales.

3.1. Operador posición proyectado, $\hat{\mathbf{r}}_{\alpha}$

Inicialmente, vamos a definir el operador de posición proyectado a una o más bandas de energía. El operador de proyección sobre el subespacio, α , referente a los momentos cristalinos de dichas bandas se define como:

$$\hat{P}_{\alpha} \equiv \sum_{n\mathbf{k} \in \alpha} |n\mathbf{k}\rangle \langle n\mathbf{k}| \quad (21)$$

Como todo operador de proyección, \hat{P}_{α} es autoadjunto ($\hat{P}_{\alpha}^{\dagger} = \hat{P}_{\alpha}$), por lo que se define el operador posición proyectado sobre el subespacio, α , como:

$$\hat{\mathbf{r}}_{\alpha} \equiv \hat{P}_{\alpha} \hat{\mathbf{r}} \hat{P}_{\alpha} \quad (22)$$

Por lo que definimos las funciones de Wannier de dicho subespacio, $|j\mathbf{R}, \alpha\rangle$, como los estados propios del operador de posición proyectado, $\hat{\mathbf{r}}_{\alpha}$, con su valor propio asociado, \mathbf{R} . No se debe confundir con un vector de la red de Bravais ya que, en general, no van a coincidir. Utilizamos el subíndice j para distinguir entre los distintos conjuntos de funciones de Wannier, dentro de los cuales se puede pasar de cualquier función a otra mediante una traslación de red. Habrá tantos posibles valores de j como ramas electrónicas haya en α .

$$\hat{\mathbf{r}}_{\alpha} |j\mathbf{R}, \alpha\rangle = \mathbf{R} |j\mathbf{R}, \alpha\rangle \quad (23)$$

Este es el procedimiento que propuso Kivelson [5], pero tan solo lo desarrolló para cristales unidimensionales. Nosotros vamos a extender el método para el caso tridimensional, con todas las dificultades y cambios que ello conlleva.

3.2. Cumplimiento de las propiedades de las funciones de Wannier

Vamos a demostrar que las funciones obtenidas mediante nuestro método alternativo cumplen las mismas propiedades que las funciones de Wannier (Ec. 19) y que, además, se encuentran máximamente localizadas.

3.2.1. Base completa de estados

Suponiendo que podemos resolver el problema de autovalores y autoestados del operador posición proyectado en tres dimensiones, los estados propios asociados a valores propios distintos son ortogonales entre sí. Además, como las ondas de Bloch son también ortogonales, los estados propios de operadores de posición proyectados a subespacios distintos también han de serlo. La normalización de dichos estados, por otro lado, es una tarea sencilla. Obtenemos directamente que:

$$\langle j'\mathbf{R}', \beta | j\mathbf{R}, \alpha \rangle = \delta_{j\mathbf{R}j'\mathbf{R}'} \delta_{\alpha\beta} \quad (24)$$

Además, sabemos que las ondas de Bloch generan un conjunto completo de funciones. Como las funciones de Wannier se generan a partir de las ondas de Bloch mediante una transformación unitaria, es evidente que:

$$\sum_{\alpha, j\mathbf{R}} |j\mathbf{R}, \alpha\rangle \langle j\mathbf{R}, \alpha| = \sum_{\alpha} \hat{P}_{\alpha} = \mathbb{I} \quad (25)$$

Ambos resultados, (Ec. 24) y (Ec. 25), indican que el conjunto de funciones de Wannier así formadas generan una base completa de funciones.

3.2.2. Traslaciones de red

Al aplicar una traslación de red, \mathbf{R}' , a una función de Wannier (Ec. 19),

$$\psi_{n\mathbf{R}}(\mathbf{r} + \mathbf{R}') \equiv \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{R}+\mathbf{R}')} \phi_{n\mathbf{k}}(\mathbf{r}) e^{i\theta_{\mathbf{k}}} \equiv \psi_{n\mathbf{R}+\mathbf{R}'}(\mathbf{r}), \quad (26)$$

se obtiene otra función de Wannier. Esta propiedad nos indica que, conocida una función de Wannier, no tenemos más que trasladarla por todos los vectores de la red de Bravais para obtener otras. En un subespacio, α , como el que nos ocupa, habrá tantas funciones de Wannier no equivalentes como ramas electrónicas lo compongan.

Llegados a este punto, conviene detenerse a aclarar un vacío en la notación de Dirac: la aplicación de los operadores. Dada la naturaleza de esta notación, en parte debida a que usualmente trabajamos con operadores autoadjuntos, al generar un sandwich no queda

claro a qué lado del producto escalar se aplica el operador. Para intentar superar esta falta de precisión, definimos:

$$(\phi, \hat{A}\psi) \equiv \langle \phi | \hat{A} | \psi \rangle, \quad (\hat{A}\phi, \psi) \equiv \left(\langle \phi | \hat{A} \right) | \psi \rangle \quad (27)$$

Aplicamos esto al estudio de las relaciones de conmutación entre el operador de proyección, \hat{P}_α , y el de traslaciones de red, $\hat{T}_{\mathbf{R}}$. Para ello, es útil tener en cuenta que:

$$\hat{T}_{\mathbf{R}} |n\mathbf{k}\rangle = e^{i\mathbf{k}\cdot\mathbf{R}} |n\mathbf{k}\rangle, \quad \left(\langle n\mathbf{k} | \hat{T}_{\mathbf{R}}^\dagger \right) = e^{i\mathbf{k}\cdot\mathbf{R}} \langle n\mathbf{k} | \quad (28)$$

Con estas relaciones, se obtiene que:

$$\hat{T}_{\mathbf{R}} \hat{P}_\alpha = \hat{T}_{\mathbf{R}} \sum_{n\mathbf{k} \in \alpha} |n\mathbf{k}\rangle \langle n\mathbf{k}| = \sum_{n\mathbf{k} \in \alpha} |n\mathbf{k}\rangle \left(\langle n\mathbf{k} | \hat{T}_{\mathbf{R}}^\dagger \right) = \sum_{n\mathbf{k} \in \alpha} |n\mathbf{k}\rangle \langle n\mathbf{k}| \hat{T}_{\mathbf{R}} = \hat{P}_\alpha \hat{T}_{\mathbf{R}} \quad (29)$$

Por lo que ambos operadores conmutan:

$$\boxed{[\hat{P}_\alpha, \hat{T}_{\mathbf{R}}] = 0} \quad (30)$$

Veamos qué sucede con el operador posición proyectado al subespacio:

$$\hat{T}_{\mathbf{R}}^\dagger \hat{\mathbf{r}}_\alpha \hat{T}_{\mathbf{R}} = \hat{P}_\alpha \hat{T}_{\mathbf{R}}^\dagger \hat{\mathbf{r}} \hat{T}_{\mathbf{R}} \hat{P}_\alpha = \hat{P}_\alpha (\hat{\mathbf{r}} + \mathbf{R}) \hat{P}_\alpha = \hat{\mathbf{r}}_\alpha + \mathbf{R} \hat{P}_\alpha \quad (31)$$

De donde se obtiene la siguiente identidad:

$$\boxed{\hat{\mathbf{r}}_\alpha \hat{T}_{\mathbf{R}} = \hat{T}_{\mathbf{R}} (\hat{\mathbf{r}}_\alpha + \mathbf{R} \hat{P}_\alpha)} \quad (32)$$

Teniendo esto en cuenta, vamos a estudiar los efectos de trasladar una función propia de $\hat{\mathbf{r}}_\alpha$ por un vector de la red de Bravais. Inicialmente se tiene que:

$$\hat{\mathbf{r}}_\alpha |j\mathbf{R}', \alpha\rangle = \mathbf{R}' |j\mathbf{R}', \alpha\rangle \quad (33)$$

Si aplicamos a dicho autoestado el operador traslación de red y volvemos a aplicarle el operador de posición proyectado, se obtiene que:

$$\hat{\mathbf{r}}_\alpha \hat{T}_{\mathbf{R}} |j\mathbf{R}', \alpha\rangle = \hat{T}_{\mathbf{R}} (\hat{\mathbf{r}}_\alpha + \mathbf{R} \hat{P}_\alpha) |j\mathbf{R}', \alpha\rangle = \hat{T}_{\mathbf{R}} (\mathbf{R}' + \mathbf{R}) |j\mathbf{R}', \alpha\rangle \quad (34)$$

Esto es:

$$\boxed{\hat{\mathbf{r}}_\alpha (\hat{T}_{\mathbf{R}} |j\mathbf{R}', \alpha\rangle) = (\mathbf{R}' + \mathbf{R}) (\hat{T}_{\mathbf{R}} |j\mathbf{R}', \alpha\rangle)} \quad (35)$$

Con lo que se demuestra que al aplicar una traslación de red, \mathbf{R} , a un estado propio del operador de posición proyectado, $|j\mathbf{R}', \alpha\rangle$, se obtiene otro estado propio de $\hat{\mathbf{r}}_\alpha$ con valor propio asociado $\mathbf{R}' + \mathbf{R}$.

3.2.3. Máxima localización

Kivelson en su artículo [5] define la desviación o segundo momento de las funciones de Wannier dentro del subespacio α , compuesto por una banda de energía aislada, como:

$$[\sigma_{\hat{\mathbf{r}}_\alpha}^2] = \frac{1}{N} \sum_{\mathbf{R}} \langle \mathbf{R}, \alpha | (\hat{\mathbf{r}}_\alpha - \mathbf{R})^2 | \mathbf{R}, \alpha \rangle \quad (36)$$

Y dado que

$$[[\hat{\mathcal{H}}, \hat{\mathbf{r}}], \hat{\mathbf{r}}] = -\frac{\hbar^2}{m_e}, \quad (37)$$

concluye que se puede obtener una cota superior de dicho segundo momento,

$$[\sigma_{\hat{\mathbf{r}}_\alpha}^2] \leq \frac{\hbar^2}{2m_e E_g}, \quad (38)$$

donde E_g es el valor del gap de energía que separa dicha banda de las demás. Este resultado asegura que las funciones de Wannier obtenidas se encuentran localizadas, pero no demuestra la máxima localización.

Para comprobar la máxima localización de los resultados, nos apoyamos en un artículo de F. Gygi *et al.* [4]. Para ello, vamos a buscar una transformación unitaria que, partiendo de las ondas de Bloch, nos dé un conjunto de funciones máximamente localizadas. Si denotamos las ondas de Bloch asociadas al subespacio α por $\phi_{n\mathbf{k}}$, la nueva base de funciones se puede definir como:

$$\psi_{j\mathbf{R}} = \sum_{n\mathbf{k} \in \alpha} x_{j\mathbf{R}n\mathbf{k}} \phi_{n\mathbf{k}} \quad (39)$$

donde $X = \{x_{j\mathbf{R}n\mathbf{k}}\}$ es una matriz unitaria restringida al subespacio α . Aunque parezca tener cuatro índices, hay que entender $(n\mathbf{k})$ y $(j\mathbf{R})$ como un único índice cada uno. Se ha de tener en cuenta que, como nos encontramos restringidos a dicho subespacio, el operador de posición usual, $\hat{\mathbf{r}}$, y el proyectado, $\hat{\mathbf{r}}_\alpha$, son totalmente equivalentes, ya que los proyectores actúan como la identidad. Por lo tanto, la desviación de una de estas ondas de Bloch se define como:

$$\sigma_{\hat{\mathbf{r}}_\alpha}^2(\phi_{n\mathbf{k}}) = \langle n\mathbf{k} | (\hat{\mathbf{r}}_\alpha - \langle \hat{\mathbf{r}}_\alpha \rangle)^2 | n\mathbf{k} \rangle = \langle n\mathbf{k} | \hat{\mathbf{r}}_\alpha^2 | n\mathbf{k} \rangle - \langle n\mathbf{k} | \hat{\mathbf{r}}_\alpha | n\mathbf{k} \rangle^2 \quad (40)$$

Si identificamos las matrices

$$a_{n\mathbf{k}n'\mathbf{k}'} \equiv \langle n\mathbf{k} | \hat{\mathbf{r}}_\alpha | n'\mathbf{k}' \rangle \quad b_{n\mathbf{k}n'\mathbf{k}'} \equiv \langle n\mathbf{k} | \hat{\mathbf{r}}_\alpha^2 | n'\mathbf{k}' \rangle, \quad (41)$$

la definición en (Ec. 40) se puede extender a todo el conjunto de las ondas de Bloch mediante la suma de todas las contribuciones:

$$\sigma_{\hat{\mathbf{r}}_\alpha}^2(\{\phi_{n\mathbf{k}}\}) = \sum_{n\mathbf{k} \in \alpha} \sigma_{\hat{\mathbf{r}}_\alpha}^2(\phi_{n\mathbf{k}}) = \text{Tr}(B) - \sum_{n\mathbf{k}} a_{n\mathbf{k}n\mathbf{k}}^2 \quad (42)$$

Esta expresión puede trasladarse al conjunto de funciones resultado de aplicar la transformación unitaria, X , tal que

$$\boxed{[\sigma_{\hat{\mathbf{r}}_\alpha}^2] \equiv \sigma_{\hat{\mathbf{r}}_\alpha}^2(\{\psi_{j\mathbf{R}}\}) = \text{Tr}(XBX^H) - \sum_{n\mathbf{k}} (XAX^H)_{n\mathbf{k}n\mathbf{k}}^2} \quad (43)$$

Por esto, el cálculo del conjunto de funciones de onda de mínima dispersión se limita a obtener aquella transformación unitaria que minimice la expresión (Ec. 43). Como bajo transformaciones unitarias la traza es invariante, tenemos que $\text{Tr}(B) = \text{Tr}(XBX^H)$, por lo que minimizar $[\sigma_{\hat{\mathbf{r}}_\alpha}^2]$ es equivalente a maximizar $\sum_{n\mathbf{k}} (XAX^H)_{n\mathbf{k}n\mathbf{k}}^2$. Al ser la norma euclídea, $\|A\|_F \equiv \sqrt{\text{Tr}(AA^H)}$, también invariante bajo transformaciones unitarias, nuestro objetivo es lograr que el valor de $\sum_{n\mathbf{k} \neq n'\mathbf{k}'} |(XAX^H)_{n\mathbf{k}n'\mathbf{k}'}|^2$ sea mínimo. Como por construcción la matriz A es hermítica, la solución de este problema se limitaría a diagonalizar la matriz o, lo que es lo mismo, resolver el problema de autoestados del operador $\hat{\mathbf{r}}_\alpha$.

En definitiva, si las funciones $\{\psi_{j\mathbf{R}}\}$ son autoestados del operador posición proyectado, $\hat{\mathbf{r}}_\alpha$, estaríamos minimizando $[\sigma_{\hat{\mathbf{r}}_\alpha}^2]$. Sin embargo, tal y como comentaremos a continuación, el problema no es tan sencillo, ya que hemos ignorado su naturaleza tridimensional.

3.3. Búsqueda de funciones de Wannier en 3D

Hasta ahora, hemos tratado a $\hat{\mathbf{r}}_\alpha$ como si fuera un único operador. Hemos deducido que, resolviendo el problema de autoestados del mismo, obtendríamos funciones de Wannier óptimamente localizadas. En el caso tridimensional, sin embargo, el operador posición proyectado está compuesto por tres operadores: uno por cada componente. Por lo tanto, si queremos diagonalizar dicho operador, tenemos que diagonalizar simultáneamente los tres operadores $\{\hat{x}_\alpha, \hat{y}_\alpha, \hat{z}_\alpha\}$.

3.3.1. No conmutabilidad

Para que tres operadores puedan diagonalizarse simultáneamente de manera exacta, es condición necesaria que éstos conmuten entre sí. Sabemos que los operadores \hat{x} , \hat{y} y \hat{z} usuales conmutan entre sí, pero no ocurre lo mismo con sus análogos proyectados. Para entenderlo, vamos a realizar un ejemplo con las componentes proyectadas x e y y una onda de Bloch, $|n\mathbf{k}\rangle$, con $n\mathbf{k} \in \alpha$.

$$\langle n'\mathbf{k}' | \hat{x}_\alpha \hat{y}_\alpha | n\mathbf{k} \rangle = \langle n'\mathbf{k}' | \hat{x} \hat{P}_\alpha \hat{y} | n\mathbf{k} \rangle \quad (44)$$

$$\langle n'\mathbf{k}' | \hat{y}_\alpha \hat{x}_\alpha | n\mathbf{k} \rangle = \langle n'\mathbf{k}' | \hat{y} \hat{P}_\alpha \hat{x} | n\mathbf{k} \rangle \quad (45)$$

El caso es que los operadores \hat{x} e \hat{y} , al ser aplicados sobre una onda de Bloch perteneciente al subespacio α , resultan en una combinación lineal arbitraria de ondas de Bloch que, en el caso más general, no van a encontrarse restringidas dentro de dicho subespacio. De este resultado, al aplicar el operador proyector, nos quedamos sólo con aquellas ondas de Bloch que pertenezcan al subespacio y, posteriormente, les aplicaremos el otro operador componente de la posición. Por lo tanto, es obvio que en este caso el orden de los factores sí altera el producto, lo que se traduce en que los operadores $\{\hat{x}_\alpha, \hat{y}_\alpha, \hat{z}_\alpha\}$ no conmutan entre sí. Otra manera de entenderlo es viendo que, en general, estos no conmutan con \hat{P}_α , ya que si aplicamos el proyector antes que el operador componente, el resultado no va a encontrarse restringido al subespacio. En cambio, si aplicamos el operador componente previamente, el operador proyector se va a encargar de que el resultado final se encuentre restringido a α . Esto hace que no podamos pasar de (Ec. 41) a (Ec. 42) mediante reglas de conmutación, por lo que ambos valores van a ser distintos (salvo que $n\mathbf{k} = n'\mathbf{k}'$).

3.3.2. Diagonalización simultánea aproximada. Ángulos de Jacobi.

El hecho de que las componentes del operador de posición proyectado no conmuten entre sí implica que no se puedan diagonalizar simultáneamente. Por consiguiente, el problema de autoestados del operador $\hat{\mathbf{r}}_\alpha$ no tiene solución exacta. Sin embargo, podemos considerar la alternativa de *diagonalizar simultáneamente de manera aproximada* un conjunto de matrices hermiticas, $A^{(c)} = \{a_{n\mathbf{k}n'\mathbf{k}'}^{(c)}\}$, donde $c = \{x, y, z\}$ y:

$$a_{n\mathbf{k}n'\mathbf{k}'}^{(c)} \equiv \langle n\mathbf{k} | \hat{c}_\alpha | n'\mathbf{k}' \rangle \quad (46)$$

Diversos autores han estudiado el problema de la *diagonalización simultánea aproximada* de un conjunto de matrices que no conmutan, es decir, obtener una transformación unitaria que convierta un grupo de m matrices máximamente diagonales considerando ciertos criterios. En nuestro método, vamos a utilizar como criterio que la suma de las componentes extradiagonales de las matrices,

$$\mathbf{off}(A) \equiv \sum_{c=x,y,z} \sum_{n\mathbf{k} \neq n'\mathbf{k}'} |a_{n\mathbf{k}n'\mathbf{k}'}^{(c)}|^2, \quad (47)$$

sea mínima. Tal y como habíamos argumentado tras (Ec. 43), este criterio es totalmente equivalente a minimizar la dispersión acumulada total de los tres operadores,

$$[\sigma_\alpha^2] = \sum_{c=x,y,z} [\sigma_{\hat{c}_\alpha}^2]. \quad (48)$$

Éste es un resultado de gran valor, ya que nos asegura que estamos obteniendo funciones máximamente localizadas. Cardoso y Souloumiac [2] propusieron un método para obtener una transformación unitaria, X , que minimiza $\mathbf{off}(A)$. Esta transformación unitaria se logra mediante la aplicación iterativa de rotaciones complejas, tal y como se explica a continuación.

Inicialmente, se inicializa la transformación unitaria, X , como una matriz compleja unidad de dimensionalidad igual al subespacio α . Es decir, si el subespacio α está compuesto por N ondas de Bloch, la matriz de cambio de base, X , pertenecerá al grupo de matrices $\mathbb{C}^{N \times N}$. A continuación, para cada par de índices $(n\mathbf{k}, n'\mathbf{k}')$ del triángulo extradiagonal inferior, se multiplica por una matriz compleja de rotación, $R(n\mathbf{k}, n'\mathbf{k}', c, s)$, que difiere de la unidad tan solo en los siguientes elementos:

$$\begin{pmatrix} 1 & 0 & \cdots & & \cdots & 0 \\ 0 & \ddots & & & & \vdots \\ \vdots & & r_{n\mathbf{k}n\mathbf{k}} & \cdots & r_{n\mathbf{k}n'\mathbf{k}'} & \\ & & \vdots & \ddots & \vdots & \\ & & & & r_{n'\mathbf{k}'n\mathbf{k}} & \cdots & r_{n'\mathbf{k}'n'\mathbf{k}'} & \vdots \\ \vdots & & & & & \ddots & 0 & \\ 0 & \cdots & & & \cdots & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & & \cdots & 0 \\ 0 & \ddots & & & & \vdots \\ \vdots & & c & \cdots & \bar{s} & \\ & & \vdots & \ddots & \vdots & \\ & & & & -s & \cdots & \bar{c} & \vdots \\ \vdots & & & & & \ddots & 0 & \\ 0 & \cdots & & & \cdots & 0 & 1 \end{pmatrix} \quad (49)$$

donde $c, s \in \mathbb{C}$ y $|c|^2 + |s|^2 = 1$. Además, los autores fueron capaces de dar expresiones explícitas tanto para s como para c :

$$\begin{aligned}
c &= \sqrt{\frac{x+r}{2r}} \\
s &= \frac{y-iz}{\sqrt{2r(x+r)}} \\
r &= \sqrt{x^2 + y^2 + z^2}
\end{aligned} \tag{50}$$

donde $[x, y, z]^T$ es el autovector correspondiente al mayor autovalor de la matriz real y simétrica

$$G = \text{Re} \left(\sum_{k=x,y,z} h_{n\mathbf{k}n'\mathbf{k}'}^H(A^{(k)}) h_{n\mathbf{k}n'\mathbf{k}'}(A^{(k)}) \right), \tag{51}$$

siendo

$$h_{n\mathbf{k}n'\mathbf{k}'}(A) \equiv [a_{n\mathbf{k}n\mathbf{k}} - a_{n'\mathbf{k}'n'\mathbf{k}'}, a_{n\mathbf{k}n'\mathbf{k}'} + a_{n'\mathbf{k}'n\mathbf{k}}, i(a_{n'\mathbf{k}'n\mathbf{k}} - a_{n\mathbf{k}n'\mathbf{k}'})] \tag{52}$$

Este autovector se elige de manera que la componente x sea siempre positiva. Esto asegura que las rotaciones se efectúen siempre en el mismo sentido, haciendo que el proceso converja con mayor rapidez. En caso de que el autovalor más alto sea degenerado, se puede elegir cualquiera de sus autovectores, siempre que cumpla que $x \geq 0$. Este proceso se repite hasta que la suma de (Ec. 47) deje de decrecer (dentro de una tolerancia dada).

Utilizando el algoritmo de Cardoso-Souloumiac, hemos desarrollado un método que calcula un conjunto de funciones máximamente localizadas en el espacio. Estas funciones se generan a través de rotaciones sobre el subespacio α . Dado que las ondas de Bloch son ortonormales entre sí y las rotaciones conservan el producto escalar, las funciones obtenidas mediante este método también serán ortonormales. Sin embargo, como estas funciones no son estados propios exactos del operador $\hat{\mathbf{r}}_\alpha$, no podemos asegurar que al aplicar una traslación de red a una de las funciones obtenidas obtengamos otra función del conjunto.

4. Método computacional y resultados

En esta sección se va a describir el proceso computacional para resolver el problema planteado. Primero, se introducirán los parámetros de la estructura considerada como ejemplo y se calcularán sus ondas de Bloch y el diagrama de bandas usando *Quantum Espresso*. Seguidamente, se obtendrán las FWOL para dicha estructura utilizando *wannier90*. Finalmente, se explicarán los algoritmos propuestos tanto para el método de Kivelson (unidimensional), como para el método tridimensional que hemos desarrollado. Los resultados obtenidos de estos dos programas se compararán con los obtenidos mediante *wannier90*.

4.1. Parámetros del problema

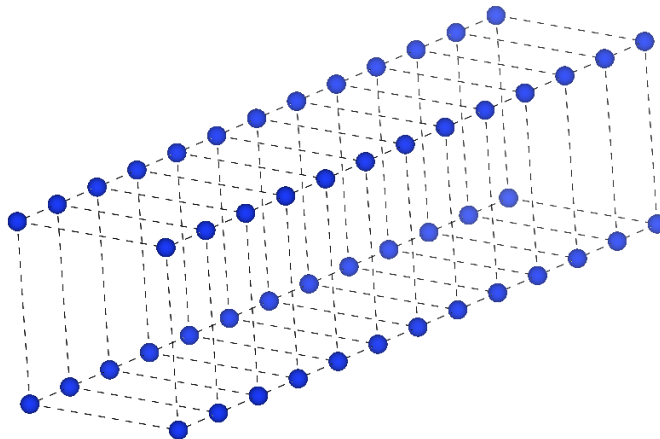


Figura 4: Representación gráfica de la estructura de silicio propuesta. Los parámetros de la celda primitiva son $6 \times 18 \times 18$ u.a.

Dado que iniciamos el proyecto siguiendo el método de Kivelson [5], diseñamos un sistema de una sola dimensión y obtenemos sus ondas de Bloch utilizando *Quantum Espresso*. Para obtener la mayor localización posible en las funciones de Wannier, tenemos que obtener una banda aislada por un gap de energía que la separe de las demás sobre la cual vamos a trabajar. Esta propiedad es análoga al caso del *enlace fuerte*, donde al ir disminuyendo la distancia interatómica, el gap de energía también lo hace (*Figura 5*). Acercar los átomos entre sí es equivalente a deslocalizar el electrón, ya que este se transportará con mayor facilidad.

La estructura considerada tiene los siguientes parámetros:

- Una celda primitiva del tipo tetragonal (ver *Figura 4*) con un átomo de silicio en el origen y unas dimensiones de:

$$\text{dim. cell.} = (6 \times 18 \times 18) \text{ u.a.} = (3,175 \times 9,525 \times 9,525) \text{ \AA} \quad (53)$$

donde u.a. son *unidades atómicas* o, lo que es lo mismo, *radios de Bohr*.

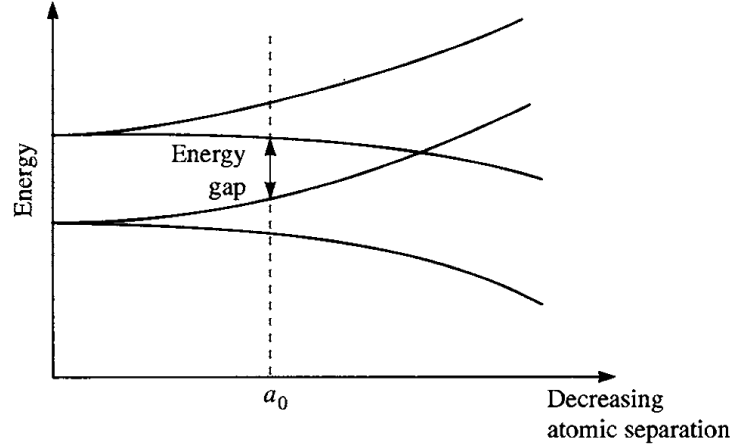


Figura 5: Al disminuir la distancia interatómica las bandas de energía asociadas a los orbitales atómicos se expanden más y más. Esta expansión se traduce en una disminución del gap de energía. Figura tomada de [8].

- Hemos considerado que el cristal se extienda a lo largo de la dirección x durante doce celdillas unidad. El número de vectores de Bravais es igual al número de vectores de momento cristalino, \mathbf{k} , que hay en la primera zona de Brillouin. Por ello, si llamamos N_i al número de divisiones de la zona de Brillouin en la dirección i , que es equivalente al número de vectores de la red de Bravais en dicha dirección, tenemos que:

$$N_x = 12, \quad N_y = 1, \quad N_z = 1 \quad (54)$$

- *Quantum Espresso* requiere que se definan ciertos parámetros, como la llamada *energía de corte*, que tiene relación directa con la precisión del cálculo que realiza. Este programa utiliza ondas planas para construir las ondas de Bloch del cristal, cuya energía cinética asociada es igual a:

$$\varepsilon_n(\mathbf{k}) = \frac{\hbar^2(\mathbf{k} + \mathbf{G}_n)^2}{2m_e} \quad (55)$$

Idealmente, la base de ondas planas es infinita, por lo que es computacionalmente imposible trabajar con ella en su totalidad. Por ello, se elige una *energía de corte*, E_{cutoff} , a partir de la cual la base es truncada. Esta aproximación es perfectamente válida para ondas de Bloch con energías asociadas mucho menores a la de corte, ya que la contribución de las ondas planas que hemos despreciado sería muy pequeña o nula. La energía de corte que hemos utilizado es la siguiente:

$$E_{\text{cutoff}} = 20 \text{ Ry} \approx 270 \text{ eV} \quad (56)$$

4.2. *Quantum Espresso* para obtener las ondas de Bloch

Aunque no se va a explicar en detalle el funcionamiento del programa, vamos a exponer brevemente los pasos realizados para la obtención de nuestro objetivo.

4.2.1. Cálculo autoconsistente

Inicialmente, se realiza un cálculo autoconsistente con la estructura deseada. Esta operación calcula la densidad de carga electrónica, $\rho(\mathbf{r})$, mediante un algoritmo similar al expuesto en la *Figura 2*. Este paso es esencial ya que esta densidad de carga es utilizada en los cálculos posteriores (que no van a ser autoconsistentes). De no hacerlo así, los resultados finales obtenidos podrían ser erróneos.

4.2.2. Cálculo de bandas

En nuestro método, hemos decidido construir las funciones de Wannier para el subespacio, α , de una banda de energía aislada. Los parámetros han sido elegidos con este objetivo, tal y como se aprecia en la *Figura 6*.

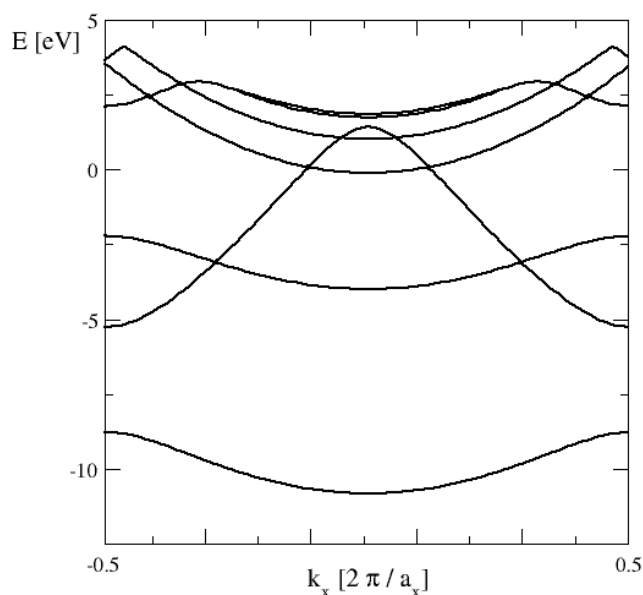


Figura 6: Diagrama de ramas electrónicas de la estructura unidimensional propuesta. Se observa que la banda de más baja energía se encuentra perfectamente aislada por un gap de unos 3eV.

Los datos representados en la *Figura 6* se han obtenido mediante una pequeña modificación del código fuente de *Quantum Espresso*. De esta manera obtenemos las energías de las distintas bandas en función de unas \mathbf{k} previamente elegidas. Este cálculo no es autoconsistente y se apoya en la densidad de carga electrónica, $\rho(\mathbf{r})$, previamente calculada.

4.2.3. Cálculo no-autoconsistente

Tras haber realizado el cálculo autoconsistente, pasamos a calcular las ondas de Bloch, $\phi_{n\mathbf{k}}(\mathbf{r})$. En nuestro caso, la banda sobre la cual vamos a trabajar (tal y como se aprecia en la *Figura 6*) es la primera, por lo que sólo nos interesarán las ondas de Bloch con $n = 1$.

Por defecto, *Quantum Espresso* guarda esta información como archivos binarios, por lo que tuvimos que modificar el código para escribir estas funciones en un archivo de texto funcional. Estos archivos no contienen las ondas de Bloch como tal, sino que contienen la función periódica, $u_{n\mathbf{k}}(\mathbf{r})$. Para su uso tendremos que multiplicarla por el valor de $e^{i\mathbf{k}\cdot\mathbf{r}}$ en

cada punto.

4.3. Referencia: Método de Marzari-Vanderbilt

Al calcular las funciones de Wannier máximamente localizadas del cristal considerado utilizando el software *wannier90* [7], obtenemos el resultado expuesto en la *Figura 7*. Este resultado lo utilizaremos como referencia para compararlo con las FWOL que obtengamos mediante el método de Kivelson y el que hemos desarrollado.

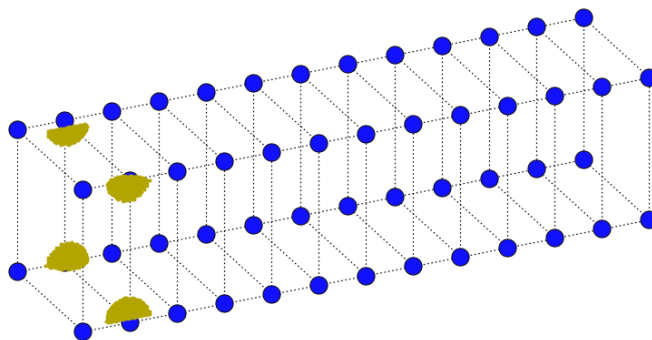


Figura 7: Zonas con valores iguales o mayores a la mitad del valor máximo del módulo de una FWOL al cuadrado. El resto de FWOL del conjunto se obtienen trasladando esta por vectores de red.

Para comparar posteriores resultados, vamos a estimar el valor de la desviación espacial típica de estas funciones. Como el valor umbral considerado en la *Figura 7* es el típicamente elegido para medir la anchura de una función, el radio de las esferas será una buena estimación de σ :

$$\sigma_{\text{M-V}} \approx 1,5 \text{ \AA} \quad (57)$$

4.4. Obtención de las funciones de Wannier óptimamente localizadas

A lo largo del trabajo hemos aplicado dos métodos. Por un lado, hemos utilizado el método de Kivelson, mediante el cual construimos las FWOL buscando los autovalores del operador \hat{x}_α , ignorando las demás componentes. Por otro lado, se han calculado mediante el método que hemos desarrollado, basado en la diagonalización simultánea de las componentes del operador posición proyectado. En cada uno de los métodos, se han obtenido resultados distintos que se expondrán a continuación.

4.4.1. Método de Kivelson

El método se basa en la idea de diagonalizar el operador posición unidimensional proyectado a una banda aislada. Para ello, hemos considerado un sistema unidimensional a lo largo del eje x , por lo que sólo diagonalizamos el operador de dicha componente; esto es, el operador \hat{x}_α . De esta manera, se van a obtener las FWOL en la dirección x , mientras que en el resto de las componentes la localización no estará asegurada.

A pesar de que se haya considerado un cristal unidimensional, no hay que olvidar que los cálculos de los sistemas de interés son siempre en tres dimensiones, así que las ondas de Bloch dependen de un vector de posición perteneciente al espacio tridimensional, $\mathbf{r} \in \mathbb{R}^3$. Es por esto que, aún suponiendo estar en una sola dimensión, las integrales que hay que realizar para calcular los elementos de matriz de $A^{(x)}$ son tridimensionales. Para explicar el método con el que se ha operado, a continuación se presenta el algoritmo utilizado de manera esquemática.

- (1) Para trabajar con una base ortonormal, se calcula la constante de normalización de cada onda de Bloch, $C(n, \mathbf{k})$, de la siguiente manera:

$$C(n, \mathbf{k}) = \sum_{\mathbf{R}} \int_{\text{cell}} d\mathbf{r} |u_{n\mathbf{k}}(\mathbf{r})|^2 = N \int_{\text{cell}} d\mathbf{r} |u_{n\mathbf{k}}(\mathbf{r})|^2 \quad (58)$$

donde $N = N_x \times N_y \times N_z$ es el número de celdillas unidad que componen el cristal. De esta manera, ya podemos construir las ondas de Bloch normalizadas a partir de las funciones periódicas que calcula *Quantum Espresso*:

$$\langle \mathbf{r} | n\mathbf{k} \rangle \equiv \phi_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{C(n, \mathbf{k})}} e^{i\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}) \quad (59)$$

- (2) Se calcula la matriz hermítica compleja, $A^{(x)} \in \mathbb{C}^{N \times N}$, definida como:

$$a_{n\mathbf{k}n'\mathbf{k}'}^{(x)} = \langle n\mathbf{k} | \hat{x}_\alpha | n'\mathbf{k}' \rangle = \langle n\mathbf{k} | \hat{x} | n'\mathbf{k}' \rangle, \quad (60)$$

Como $A^{(x)}$ es hermítica ($a_{ij}^{(x)} = \bar{a}_{ji}^{(x)}$), solo hay que realizar la mitad de las integrales extradiagonales. La expresión explícita de dicha integral es la siguiente:

$$\langle n\mathbf{k} | x | n'\mathbf{k}' \rangle = \sum_{\mathbf{R}} e^{i(\mathbf{k}' - \mathbf{k})\cdot\mathbf{R}} \left[\int_{\text{cell}} d\mathbf{r} x e^{i(\mathbf{k}' - \mathbf{k})\cdot\mathbf{r}} \bar{u}_{n\mathbf{k}}(\mathbf{r}) u_{n'\mathbf{k}'}(\mathbf{r}) + R_x \int_{\text{cell}} d\mathbf{r} e^{i(\mathbf{k}' - \mathbf{k})\cdot\mathbf{r}} \bar{u}_{n\mathbf{k}}(\mathbf{r}) u_{n'\mathbf{k}'}(\mathbf{r}) \right] \quad (61)$$

- (3) Se diagonaliza dicha matriz para obtener sus autovectores¹. Siendo $c_{j\mathbf{R}n\mathbf{k}}$ los coeficientes de los vectores propios de la matriz $A^{(x)}$, las FWOL en la dirección x , $\psi_{j\mathbf{R}}(\mathbf{r})$, se definen de la siguiente manera:

$$\psi_{j\mathbf{R}}(\mathbf{r}) = \sum_{n\mathbf{k} \in \alpha} c_{j\mathbf{R}n\mathbf{k}} \phi_{n\mathbf{k}}(\mathbf{r}) \quad (62)$$

¹Para diagonalizar la matriz se ha utilizado la subrutina *zhbev* de los paquetes *LAPACK* [1].

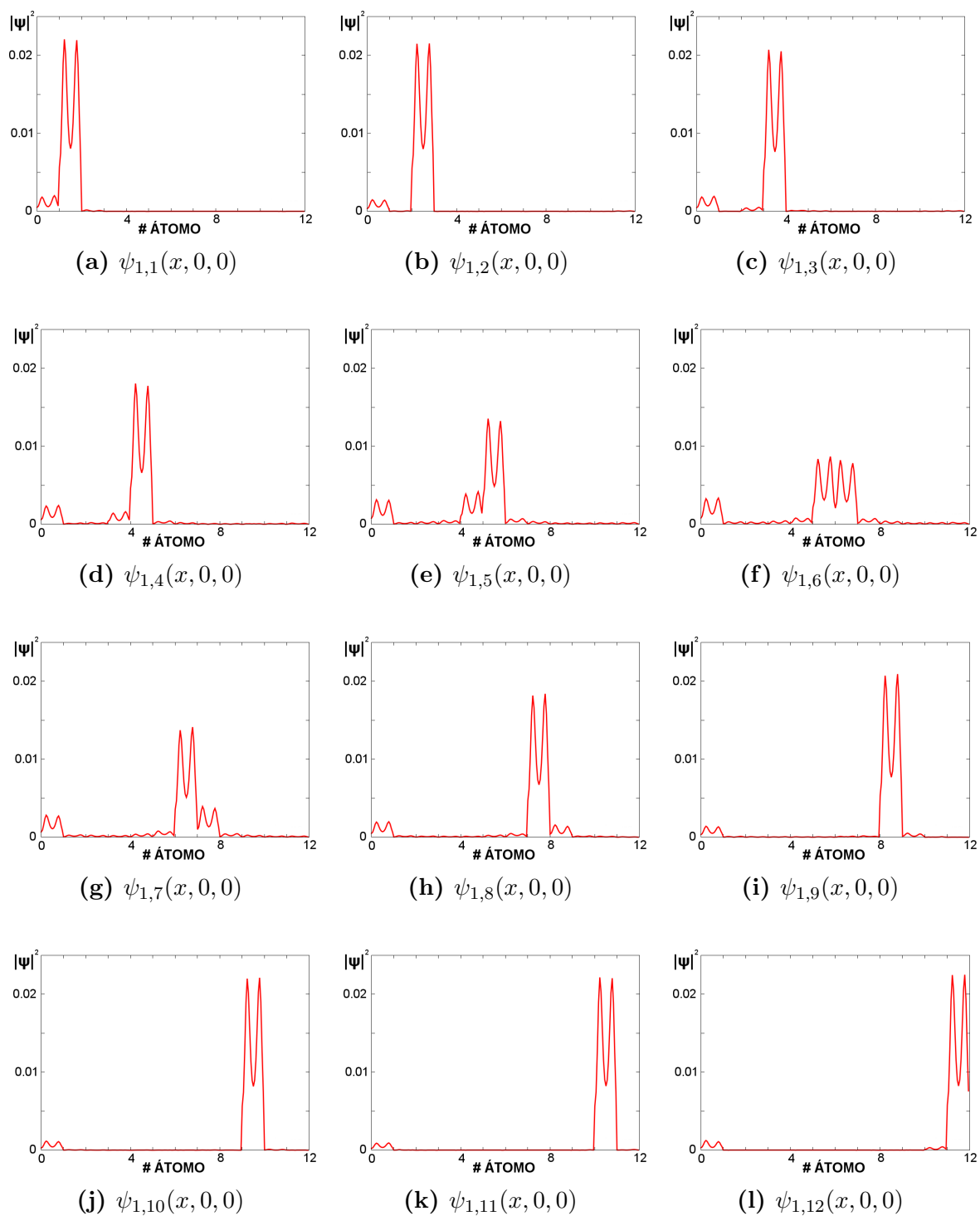


Figura 8: Funciones de Wannier calculadas considerando la diagonalización de la componente \hat{x}_α para un cristal con $N_x = 12$ y representadas sobre la línea que une los átomos de silicio.

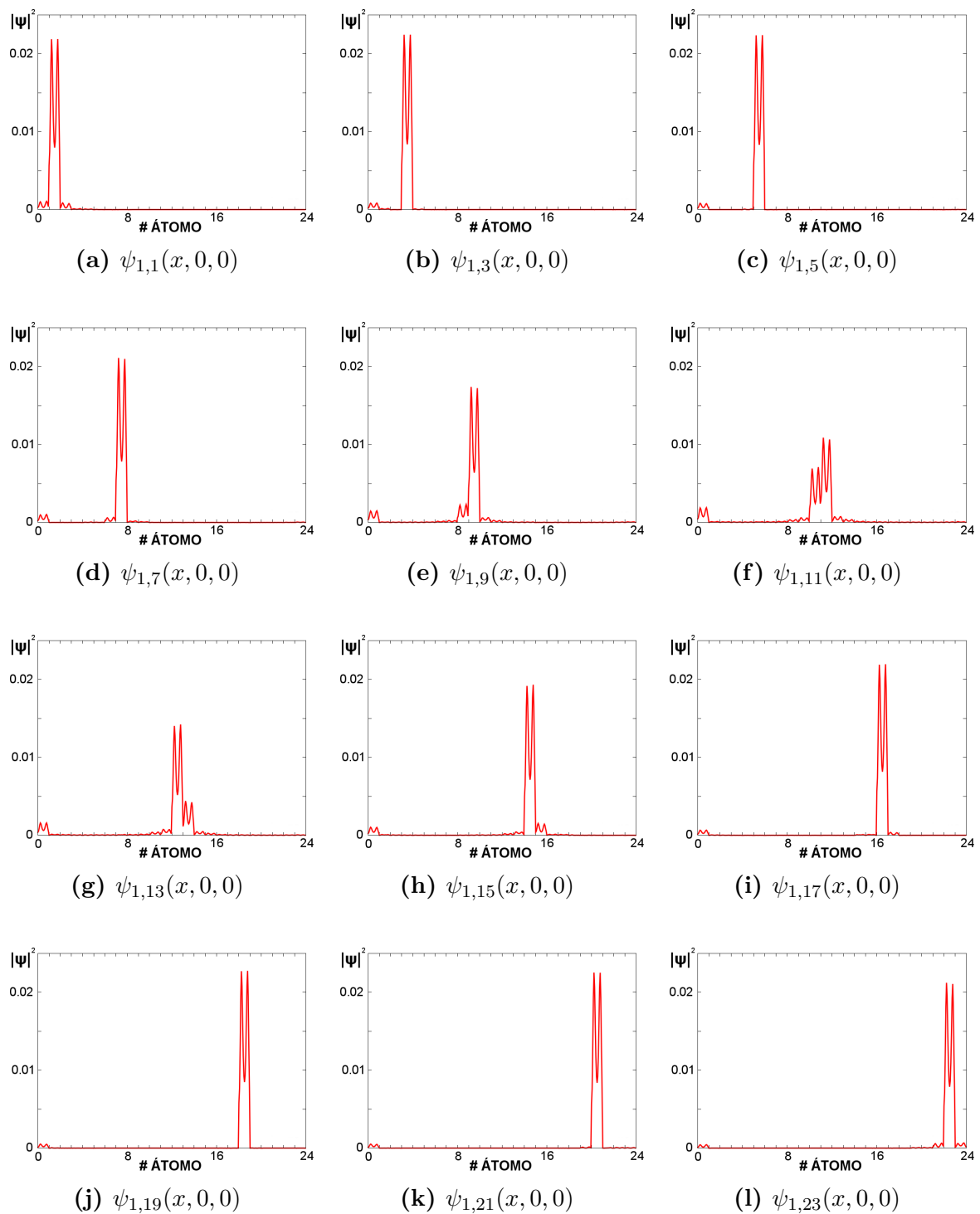


Figura 9: Funciones de Wannier calculadas considerando la diagonalización de la componente \hat{x}_α para un cristal con $N_x = 24$ y representadas sobre la línea que une los átomos de silicio.

A modo test, se han obtenido estas funciones de Wannier en dos casos distintos: con un cristal compuesto de 12 celdillas unidad dispuestas en la dirección x (*Figura 8*) y con el mismo cristal, pero esta vez compuesto por 24 celdillas (*Figura 9*). Estos gráficos representan el módulo al cuadrado de las funciones en los puntos sobre la recta que une la cadena de átomos de silicio, esto es, sobre la recta que cumple que $(x, y, z) = (x, 0, 0)$.

Al comparar las gráficas de la *Figura 8* con los de la *Figura 9* vemos que, aunque se doble la longitud de la cadena, las “funciones de Wannier” siguen siendo las mismas o muy parecidas. Esto nos indica que los resultados obtenidos a través de este método son consistentes.

Sin embargo, hay ciertas anomalías que no se han logrado explicar. En todos los casos se observa una pequeña perturbación cerca del origen de coordenadas; parece un error sistemático, pero tras una exhaustiva revisión del código fuente no se ha encontrado ningún error. Además, las funciones de Wannier asociadas a los átomos centrales del cristal son distintas; esto no debería ser así ya que, debido a las *condiciones de contorno de Born-Von Karman*, la elección del primer átomo de la red es totalmente arbitraria.

Se ha comparado el valor de la desviación típica estimada de estas gráficas con la obtenida mediante el método de Marzari-Vanderbilt y la obtenida según la fórmula que propuso Kivelson (Ec. 38). Teniendo en cuenta todas ellas, se estima que la desviación típica se encuentra entre los siguientes valores:

$$1,5 \text{ \AA} \lesssim \sigma_{\text{Kiv}} \lesssim 3,0 \text{ \AA} \quad (63)$$

En resumen, exceptuando en los resultados asociados a los átomos centrales, la localización concuerda con la obtenida mediante *wannier90* ($\sigma_{\text{M-V}} \approx 1,5 \text{ \AA}$). Por otra parte, el cálculo propuesto por Kivelson (Ec. 38) para un valor de $E_g \approx 3eV$ (*Figura 6*), devuelve un valor de:

$$\sigma \lesssim 1,2 \text{ \AA} \quad (64)$$

Valor que, aún no siendo compatible, es cercano a los resultados.

4.4.2. Método de diagonalización simultánea

El método de Kivelson es aplicable en cristales unidimensionales. Sin embargo, el autor no se percató del problema que surge al aplicar su método a cristales tridimensionales. En esos casos, el operador posición proyectado se compone de tres operadores, uno por componente. Como estos operadores, en general, no conmutan entre sí, el problema de diagonalizar simultáneamente todos ellos no tiene solución exacta. Por ello, hemos desarrollado un método basado en la diagonalización simultánea aproximada de las tres componentes de $\hat{\mathbf{r}}_\alpha$ siguiendo los pasos propuestos por Cardoso-Souloumiac [2]. Aunque este método es tridimensional, se ha considerado el mismo cristal para poder comparar los resultados con los del método de Kivelson. A continuación, se detalla un gui3n con el algoritmo que hemos utilizado en este método, paso a paso:

- (1) Se leen los ficheros obtenidos de *Quantum Espresso* con las funciones periódicas, $u_{n\mathbf{k}}(\mathbf{r})$, y se construye la base de ondas de Bloch normalizadas de la misma manera que se realizó en (Ec. 59).

- (2) Se computan las tres matrices hermíticas complejas, $A^{(x)}$, $A^{(y)}$ y $A^{(z)}$, todas pertenecientes al grupo de matrices $\mathbb{C}^{N \times N}$, donde N es la dimensión del subespacio α :

$$\begin{aligned} a_{n\mathbf{k}n'\mathbf{k}'}^{(x)} &= \langle n\mathbf{k}|x|n'\mathbf{k}' \rangle \\ a_{n\mathbf{k}n'\mathbf{k}'}^{(y)} &= \langle n\mathbf{k}|y|n'\mathbf{k}' \rangle \\ a_{n\mathbf{k}n'\mathbf{k}'}^{(z)} &= \langle n\mathbf{k}|z|n'\mathbf{k}' \rangle \end{aligned}$$

- (3) Se inicializa a la unidad la matriz que contendrá la transformación unitaria de cambio de base:

$$X = \{x_{j\mathbf{R}n\mathbf{k}}\} = \mathbb{I}$$

- (4) (*Ciclo*) Para cada par de índices $(n\mathbf{k}, n'\mathbf{k}')$ del triángulo extradiagonal superior de las matrices $N \times N$:

- (a) Se forma la matriz 3×3 ,

$$G = \text{Re} \left(\sum_{k=x,y,z} h_{n\mathbf{k}n'\mathbf{k}'}^H(A^{(k)}) h_{n\mathbf{k}n'\mathbf{k}'}(A^{(k)}) \right),$$

donde

$$h_{n\mathbf{k}n'\mathbf{k}'}(A) \equiv [a_{n\mathbf{k}n\mathbf{k}} - a_{n'\mathbf{k}'n'\mathbf{k}'}, a_{n\mathbf{k}n'\mathbf{k}'} + a_{n'\mathbf{k}'n\mathbf{k}}, i(a_{n'\mathbf{k}'n\mathbf{k}} - a_{n\mathbf{k}n'\mathbf{k}'})].$$

- (b) Se calculan los valores y vectores propios de G^2 . Entre los resultados, se toma el vector propio, $[x, y, z]^T$, asociado al valor propio más grande. Elegimos el signo de dicho vector de tal manera que $x \geq 0$. Esta elección es importante, ya que, por un lado, al mantener el signo de la componente las rotaciones serán siempre en el mismo sentido y, al ser este signo positivo, se evitarán posibles problemas de divisiones entre cero al calcular la cantidad, s , de (Ec. 50).
- (c) Se calcula la matriz de rotación compleja $N \times N$, $R(n\mathbf{k}, n'\mathbf{k}', c, s)$, que tan solo difiere de la matriz identidad en las siguientes componentes:

$$\begin{pmatrix} 1 & 0 & \cdots & & \cdots & 0 \\ 0 & \ddots & & & & \vdots \\ \vdots & & r_{n\mathbf{k}n\mathbf{k}} & \cdots & r_{n\mathbf{k}n'\mathbf{k}'} & \\ & & \vdots & \ddots & \vdots & \\ & & r_{n'\mathbf{k}'n\mathbf{k}} & \cdots & r_{n'\mathbf{k}'n'\mathbf{k}'} & \vdots \\ \vdots & & & & \ddots & 0 \\ 0 & \cdots & & \cdots & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & & \cdots & 0 \\ 0 & \ddots & & & & \vdots \\ \vdots & & c & \cdots & \bar{s} & \\ & & \vdots & \ddots & \vdots & \\ & & -s & \cdots & \bar{c} & \vdots \\ \vdots & & & & \ddots & 0 \\ 0 & \cdots & & \cdots & 0 & 1 \end{pmatrix}$$

donde:

$$\begin{aligned} c &= \sqrt{\frac{x+r}{2r}} \\ s &= \frac{y-iz}{\sqrt{2r(x+r)}} \\ r &= \sqrt{x^2 + y^2 + z^2} \end{aligned}$$

²Para esto se ha utilizado la subrutina *dsyev* de los paquetes de álgebra lineal *LAPACK* [1].

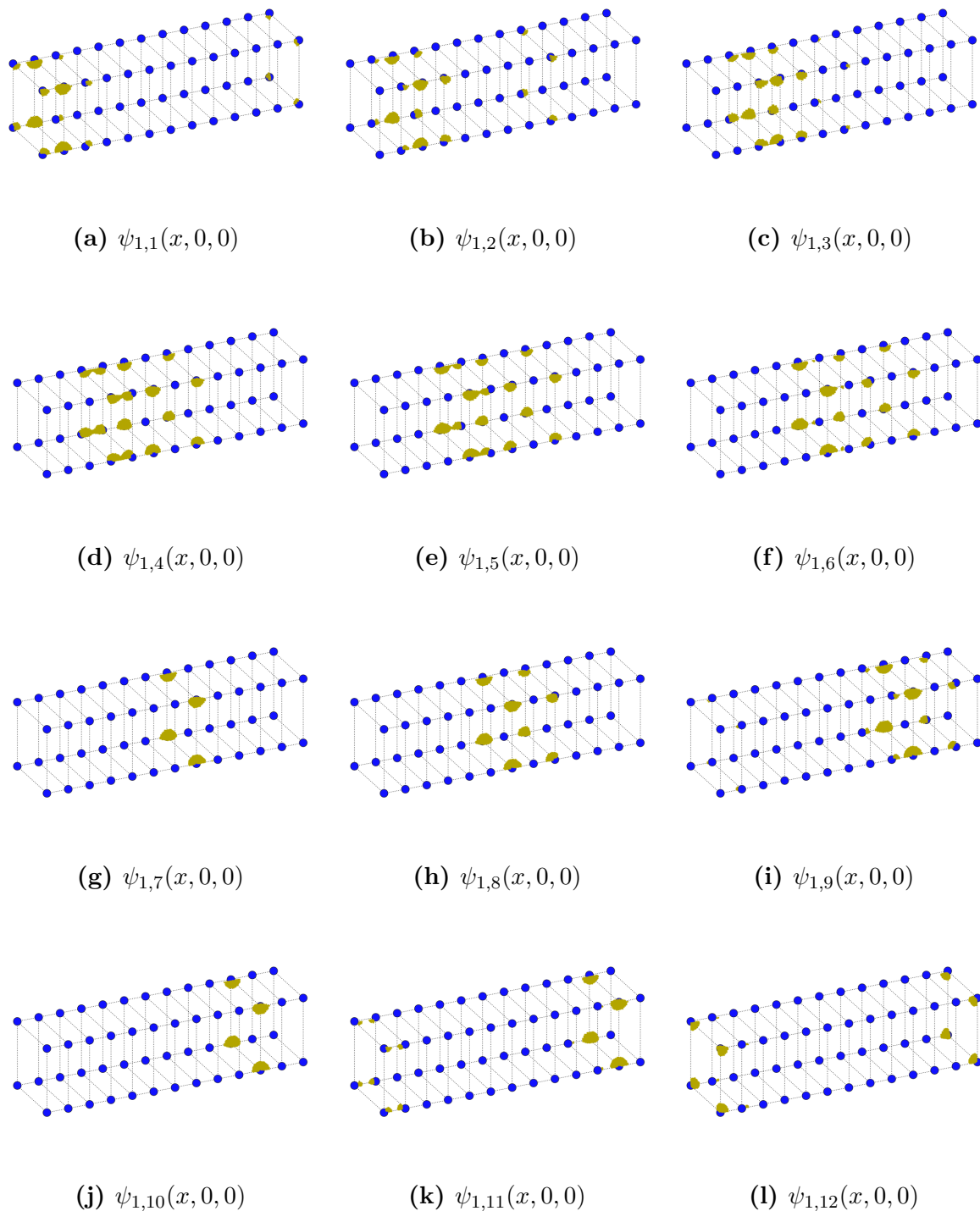


Figura 10: *Funciones de Wannier calculadas segun el método tridimensional para un cristal con $N_x = 12$. En la representación se exponen aquellos puntos del espacio donde el módulo la cuadrado de la función es mayor a la mitad de su valor máximo.*

(d) Se acumula la rotación en la transformación ortogonal, $X := RX$.

(5) Se calcula

$$\mathbf{off}(A) \equiv \sum_{c=x,y,z} \sum_{n\mathbf{k} \neq n'\mathbf{k}'} |a_{n\mathbf{k}n'\mathbf{k}'}^{(c)}|^2$$

como indicador del estado de diagonalización.

(6) Si durante el último ciclo $\mathbf{off}(A)$ decrece en un valor mayor a la tolerancia, que en nuestro caso ha sido elegida como $\tau = 10^{-7}$, se vuelve al paso 4.

(7) Se definen las FWOL, $\{\psi_{j\mathbf{R}}(\mathbf{r})\}$, de la siguiente manera:

$$\psi_{j\mathbf{R}}(\mathbf{r}) = \sum_{n\mathbf{k} \in \alpha} x_{j\mathbf{R}n\mathbf{k}} \phi_{n\mathbf{k}}(\mathbf{r}) \quad (65)$$

Al aplicar este método al cristal considerado como ejemplo, se han obtenido los resultados expuestos en la *Figura 10*, de los cuales se pueden obtener ciertas conclusiones.

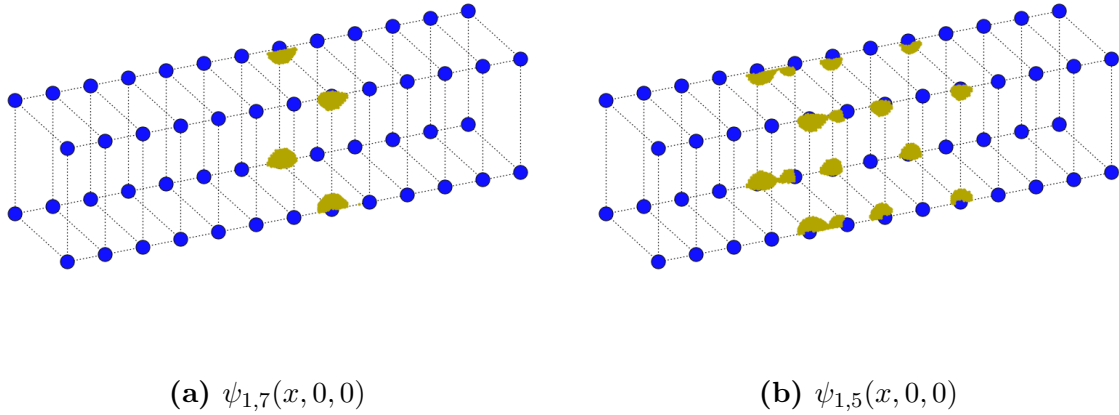


Figura 11: *Dos de las funciones de Wannier tomadas de la Figura 10 para analizar la localización de estas. Se han tomado (a) de las que más localizadas se encuentran y (b) de las que menos.*

Aunque entre las funciones representadas, algunas son similares entre sí, no estamos ante una única función trasladada por los vectores de red. Ya habíamos adelantado este comportamiento al comprobar que las funciones calculadas no son autoestados del operador $\hat{\mathbf{r}}_\alpha$. Esto indica que estas funciones no cumplen la definición original de Wannier [9] expuesta en (Ec. 2). Este incumplimiento se verifica fácilmente al comprobar que no todos los elementos de la matriz unitaria, X , tienen el mismo módulo. Para este método, tendríamos que utilizar una definición menos restrictiva:

$$\psi_{j\mathbf{R}}(\mathbf{r}) = \sum_{n\mathbf{k} \in \alpha} e^{i\mathbf{k} \cdot \mathbf{R}} \phi_{n\mathbf{k}}(\mathbf{r}) c_{j\mathbf{R}n\mathbf{k}} \quad (66)$$

donde $\sum_{n\mathbf{k} \in \alpha} |c_{j\mathbf{R}n\mathbf{k}}|^2 = 1$, ya que los resultados están normalizadas. Esta definición, sin embargo, es tan poco restrictiva que cualquier función normalizada, producto de una combinación lineal de ondas de Bloch pertenecientes a α , la cumple. Justificamos llamar

a estos resultados *funciones de Wannier* por dos razones: se encuentran localizadas en el espacio y podemos asociar cada una de ellas a una celdilla unidad distinta.

Para analizar la localización de las FWOL obtenidas, en la *Figura 11* se han aislado dos de las gráficas de la *Figura 10*. Se observa que la localización a lo largo de las direcciones y y z es excelente, pero para la dirección x esta varía de una función a otra. En ciertas funciones, como por ejemplo la expuesta en la *Figura 11(a)*, la localización en dicha dirección es muy alta (de hecho, es la más alta), mientras que en otros casos (*Figura 11(b)*, la función menos localizada del conjunto) la localización en x ya no es tan buena. Se calcularán los valores de $\sigma \equiv \sigma_{\mathbf{r}}$ para las dos funciones representadas en la *Figura 11*. De esta manera estimaremos el rango de localizaciones en las que se encuentran nuestras FWOL. Tenemos que:

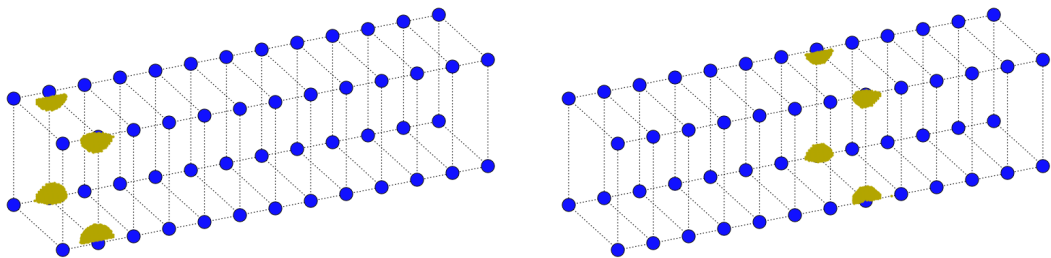
$$\sigma^2 \equiv \sigma_{\mathbf{r}}^2 = \sigma_x^2 + \sigma_y^2 + \sigma_z^2 \quad (67)$$

Del primero de los gráficos se estima de manera inmediata que $\sigma(a) \approx 1,5 \text{ \AA}$. Además, al ser una localización esférica, las desviaciones típicas de cada componente han de ser iguales. Entonces, podemos dividir por un factor $\sqrt{3}$ las anchuras máximas en cada dimensión para estimar los valores de σ_i en dichas direcciones. De este modo se obtiene que:

$$\left\{ \begin{array}{l} \sigma(a) \approx 1,5 \text{ \AA} \\ \sigma_x(a) \approx 0,87 \text{ \AA} \\ \sigma_y(a) \approx 0,87 \text{ \AA} \\ \sigma_z(a) \approx 0,87 \text{ \AA} \end{array} \right\}, \quad \left\{ \begin{array}{l} \sigma(b) \approx 4,7 \text{ \AA} \\ \sigma_x(b) \approx 4,5 \text{ \AA} \\ \sigma_y(b) \approx 0,87 \text{ \AA} \\ \sigma_z(b) \approx 0,87 \text{ \AA} \end{array} \right\} \quad (68)$$

Por lo que se estima que las desviaciones de todas las funciones obtenidas mediante el método que hemos desarrollado se encuentran entre los siguientes valores:

$$1,5 \text{ \AA} \lesssim \sigma_{ds} \lesssim 4,7 \text{ \AA} \quad (69)$$



(a) *Marzari-Vanderbilt*

(b) *Este trabajo*

Figura 12: Comparativa entre el método de (a) *Marzari-Vanderbilt* y (b) el desarrollado en este trabajo.

Por lo tanto, la localización de las funciones mejor localizadas obtenidas mediante este método concuerda con la de las obtenidas mediante *wannier90*. En la *Figura 12* puede apreciarse el excelente acuerdo que hay entre ambos casos.

5. Conclusiones

Basándonos en la diagonalización simultánea de las componentes del operador \hat{r}_α , hemos desarrollado un método alternativo *wannier90* para obtener FWOL. Este método resulta ser mucho más sencillo que el de *Marzari-Vanderbilt*. Asimismo, ha demostrado ser muy eficaz para sistemas pequeños, ya que solo necesitó siete ciclos para hallar la convergencia. De todas formas, su eficiencia para sistemas más grandes está por comprobar, ya que trabajamos con matrices cuya dimensión depende de manera lineal del número de celdas unidad del cristal, N . Por ello, se espera una dependencia del tiempo de ejecución con las dimensiones del cristal del tipo $t_{\text{ejec.}} \sim N^2$.

Se han comparado las FWOL de *wannier90* con las del método de Kivelson (para cristales unidimensionales) y con las del método que hemos desarrollado. La localización de las funciones obtenidas mediante el método de Kivelson, salvo anomalías en algunos casos, es compatible con las FWOL de *wannier90*. Las funciones de Wannier obtenidas mediante nuestro método, sin embargo, tienen un inconveniente: al trasladar una de las funciones por un vector de red no obtenemos otra función del conjunto. Esto ocurre porque dichas funciones no son autoestados del operador posición proyectado. Dada esta peculiaridad, la localización de las funciones varía de unas a otras. En el mejor de los casos, los resultados concuerdan perfectamente con aquellos obtenidos mediante el método de Marzari-Vanderbilt. En el peor de los casos, aunque los resultados no concuerden, la localización del electrón sigue siendo aceptable.

En lo que a mí respecta, este trabajo ha sido útil para conocer las bases de los cálculos electrónicos a partir de primeros principios. He comprendido las bases de la teoría del funcional de densidad (DFT) y he experimentado la utilización de programas como *Quantum Espresso* y *wannier90*.

En resumen, considero que el desarrollo del trabajo ha sido satisfactorio ya que ambos objetivos, tanto demostrar teóricamente que mediante el método desarrollado se pueden calcular FWOL como implementar un programa para aplicarlo, han sido realizados.

Se puede seguir trabajando para mejorar los tiempos de ejecución paralelizando el código o desarrollando un método de diagonalización simultánea que devuelva un resultado compatible con la ecuación original de Wannier (Ec. 2). El código fuente desarrollado para el programa basado en el método de la diagonalización simultánea estará disponible en la plataforma.

Referencias

- [1] *Linear Algebra PACKage (LAPACK)*. <http://www.netlib.org/lapack/>, Accedido en 2016.
- [2] J.-F. CARDOSO AND A. SOULOUMIAC, *Jacobi angles for simultaneous diagonalization*, SIAM J. Mat. Anal. Appl., 17 (1996), pp. 161–164.
- [3] P. GIANNOZZI, S. BARONI, N. BONINI, M. CALANDRA, R. CAR, C. CAVAZZONI, D. CERESOLI, G. L. CHIAROTTI, M. COCCIONI, I. DABO, A. DAL CORSO, S. DE GIRONCOLI, S. FABRIS, G. FRATESI, R. GEBAUER, U. GERSTMANN, C. GOUGOUSSIS, A. KOKALJ, M. LAZZERI, L. MARTIN-SAMOS, N. MARZARI, F. MAURI, R. MAZZARELLO, S. PAOLINI, A. PASQUARELLO, L. PAULATTO, C. SBRACCIA, S. SCANDOLO, G. SCLAUZERO, A. P. SEITSONEN, A. SMOGUNOV, P. UMARI, AND R. M. WENTZCOVITCH, *Quantum espresso: a modular and open-source software project for quantum simulations of materials*, Journal of Physics: Condensed Matter, 21 (2009), p. 395502 (19pp).
- [4] F. GYGI, J.-L. FATTEBERT, AND E. SCHWEGLER, *Computation of maximally localized wannier functions using a simultaneous diagonalization algorithm*, Computer Physics Communications, 155 (2003), pp. 1 – 6.
- [5] S. KIVELSON, *Wannier functions in one-dimensional disordered systems: Application to fractionally charged solitons*, Phys. Rev. B, 26 (1982), pp. 4269–4277.
- [6] N. MARZARI AND D. VANDERBILT, *Maximally localized generalized wannier functions for composite energy bands*, Phys. Rev. B, 56 (1997), pp. 12847–12865.
- [7] A. A. MOSTOFI, J. R. YATES, G. PIZZI, Y.-S. LEE, I. SOUZA, D. VANDERBILT, AND N. MARZARI, *An updated version of wannier90: A tool for obtaining maximally-localised wannier functions*, Computer Physics Communications, 185 (2014), pp. 2309 – 2310.
- [8] A. P. SUTTON, *Electronic structure of materials*, Oxford Science Publications, Oxford University Press, 1993.
- [9] G. H. WANNIER, *The structure of electronic excitation levels in insulating crystals*, Phys. Rev., 52 (1937), pp. 191–197.