



Propiedades Físicas de las Cápsides Virales Icosaédricas:

**Modelos de Potenciales de Interacción
y Constantes de Fuerza**

ADRIÁN BACALLADO RIVERO

Trabajo de Fin de Grado

SUPERVISADO POR:
JOSÉ MARÍA GÓMEZ LLORENTE

Agradecimientos

Me gustaría agradecer a mi tutor, José María, por aceptarme en la realización de este trabajo, y por tener paciencia y confiar en mi.

Agradecer, también, a los profesores que he tenido estos cuatro años de carrera, que sin duda, me han aportado una nueva forma de pensar y trabajar.

Me gustaría agradecer, de corazón, a todos mis compañeros de clase, por su apoyo, ofrecerme su ayuda o por amenizar esta dura carrera.

Agradecer, sobre todo, a mi familia, por la ayuda incondicional que me han dado siempre que la he necesitado.

Por último, me gustaría agradecer a mis amigos Kevin y Raúl, que sin su apoyo y consejos no podría haberlo conseguido. Agradecer, de nuevo, a Kevin por ayudarme a resolver los fallos del compilador de LaTeX que no conseguía arreglar.

Gracias, a todos.

Índice

1. Abstract	1
2. Introducción	2
2.1. Objetivo	4
3. Estructura de los virus icosaédricos	6
3.1. Modelo de Caspar y Klug	7
4. Formación de la cápside	10
4.1. Termodinámica del ensamblaje	10
4.2. Cinética del ensamblaje	12
5. Propiedades físicas de los virus	15
5.1. Electroestática	15
5.2. Propiedades mecánicas	17
6. Modelos de interacción	21
6.1. Modelos <i>All-atom</i>	21
6.2. Modelos <i>Coarse-grained</i>	22
7. Aplicación práctica	24
7.1. Potencial de interacción	24
7.2. Cálculo de las constantes de fuerza	27
8. Conclusiones	30
Referencias	31
A. Cálculo de las constantes de fuerza	34

1. Abstract

Virology is a research field that needs physics to understand the behaviour of viruses, since there are a lot mechanisms that use thermodynamics, kinetics or electrostatics. These are some of the viral properties that we are going to study and explain.

In this final degree work, we will start explaining the viruses in their fundamentals, introducing the capsids, the envelope of the viruses. We will consider only the case of the icosahedral viruses, since this geometrical form is the one that appears the most in nature. These capsids are formed by protein subunits, the capsomers. Icosahedral capsids are described by Caspar and Klug's models, since they introduce the triangular number T , a very important parameter in virology. One of the most fastinating feature is the auto-assembly of viral capsids. This is a feature that we will explain via thermodynamics and kinetics.

We will study as well the electrostatic interaction between the capsomers and the capsomers in the formed capsid through Poisson-Boltzmann's equation. Another physical feature that we will study is the mechanical properties. Viruses endure external forces in their environment and the osmotic pressure that the genome applies to the capsid. That force can be measured and studied.

The main focus of this work will be on the models that explain the interaction potential. First, we will explain the two different models that explain these interactions: *Coarse-grained* and *All-atom*. Then, we will explain our two-body interaction model, that is a *Coarse-grained* type, using trimers, a kind of triangular capsomers. Afterwards, we will introduce the variables that characterize the trimer orientation and the equilibrium conditions that fixes the privileged orientation of the trimers in order to form a capsid. Then, we will calculate the second derivative matrix of the interaction potential in order to calculate the force constants. Finally, we apply the equilibrium conditions to the matrix to obtain the force constants.

2. Introducción

Abstract. *The viruses are infectious biological entities that are formed by the capsid, the envelope of the viruses, and the genetical information, packaged in DNA or RNA inside the capsid. This capsid is made of protein subunits called capsomers. These capsomers can be pentamers, with pentagonal form or hexamers, with hexagonal form. There are different virus forms: icosahedral, cylindrical or complex. We will explain the objectives of this work in this chapter, such as explaining the interaction models that describe the two-body interaction potential or the calculus of the force constants.*

En la historia de la humanidad, los virus han sido considerados una amenaza, dado la gran cantidad de vidas humanas que se han perdido por infecciones virales. Hoy en día, la medicina ha mejorado, ayudando así a neutralizar los virus, pero a veces no es suficiente. Esta es una de las causas por las que estamos estudiando los virus y sus propiedades, para entender como trabajan y utilizar esta información a nuestro favor.

Los virus son cuerpos biológicos microscópicos de naturaleza infecciosa que necesitan una célula huésped para replicarse a si mismos. Están compuestos por el genoma y una cápside, y algunos virus poseen una bicapa lipídica. El genoma tiene la información genética del virus y puede ser ARN (ácido ribonucleico) o ADN (ácido desoxirribonucleico), y trenzado simple o doble trenzado, de manera que los virus tienen diferentes formas de empaquetar el genoma.

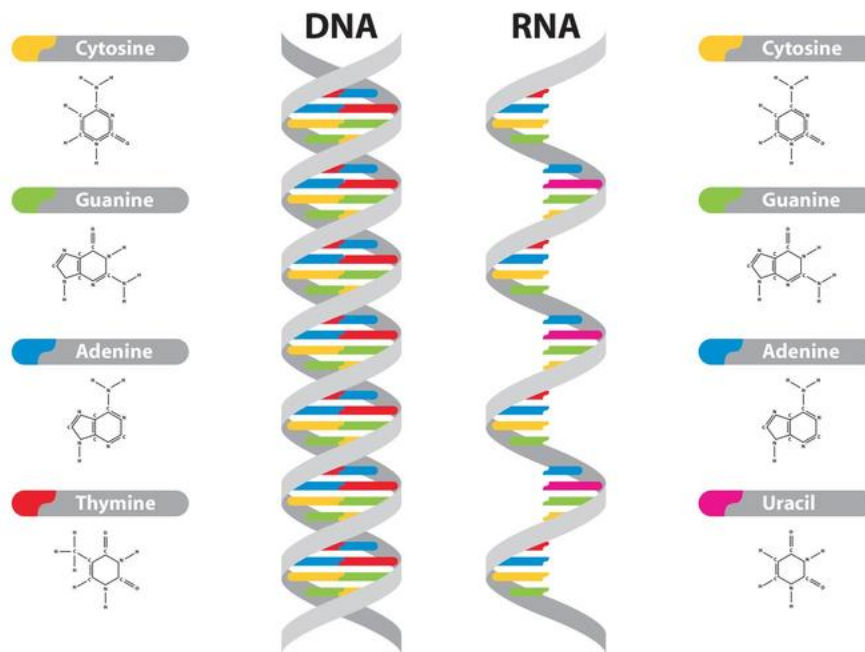


Figura 1: Moléculas de ADN y ARN

La cápside es una estructura que protege el genoma y está formado por unidades proteicas fundamentales, llamadas capsómeros. Estos capsómeros pueden estar formados por una o varias proteínas del mismo tipo, o incluso proteínas de distinto tipo. Además, los capsómeros pueden ser hexámeros, que están rodeados por seis capsómeros vecinos, es decir, tienen forma hexagonal, o pentámeros, rodeados por cinco capsómeros vecinos, es decir, tienen forma pentagonal. Las cápsides víricas pueden tener diferentes formas: icosaédrica, como el adenovirus, helicoidal, como el virus del mosaico del tabaco (TMV), o compleja, como el Bacteriófago T4 [1].

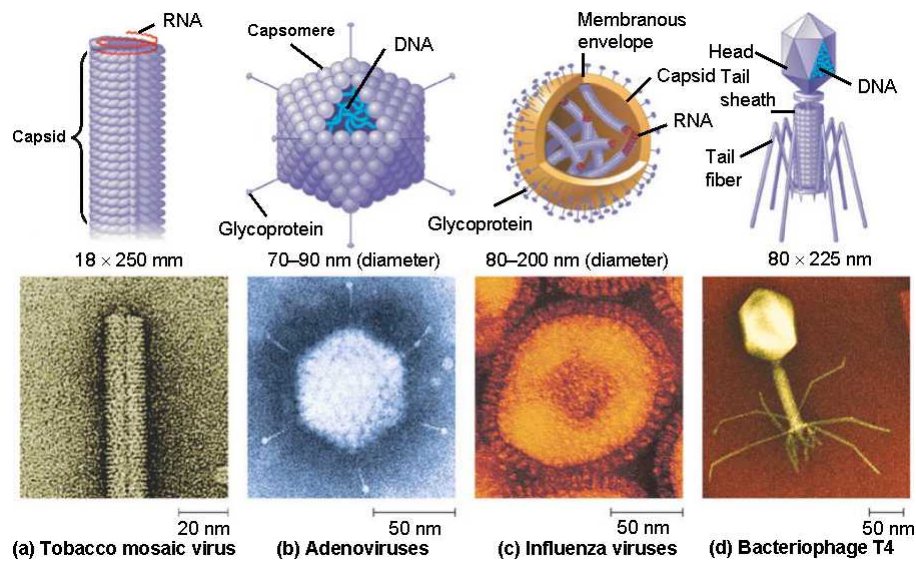


Figura 2: Ejemplos de las distintas formas de los virus: a) TMV, b) Adenovirus, c) Influenza virus, d) Bacteriófago T4.

La formación de virus requiere el montaje de la cápside, envolvimiento por una membrana, si la posee, y el empaquetamiento del genoma dentro. Los virus con bicapa lipídica, como el VIH o el Influenza (ver Figura 2), obtienen esta membrana de la célula huésped. Una de las características más sorprendentes de los virus es que muchos de los virus de trenzado simple se autoensamblan de manera espontánea alrededor del material genético. Sin embargo, los virus de doble trenzado, debido a la dureza y la alta densidad de carga, imposibilita el autoensamblaje. Por ello, muchos dsDNA (ADN de doble trenzado) virus montan una cápside proteica vacía (procápside) y un motor molecular que hidroliza ATP para inyectar el ADN en la cápside [2].

2.1. Objetivo

El objetivo principal de este trabajo consiste en explicar los distintos tipos de modelos de interacción entre capsómeros en los que se trabajan actualmente y emplearlos para hacer una primera aproximación al cálculo de las constantes de fuerza asociados a la interacción de dos capsómeros, los cuales vamos a considerar sólidos rígidos. Por ello, utilizaremos la metodología típica de la Mecánica clásica para afrontar nuestro problema.

Las constantes de fuerza son las constantes de proporcionalidad que existen entre una fuerza aplicada y un desplazamiento realizado en cada dirección que corresponde a las variables que describen el problema. Estas variables se relacionan con los grados de libertad que posee el problema. Estas constantes aparecen al obtener la matriz de segundas derivadas del potencial de interacción que se emplea con respecto a las variables que caracterizan el problema. En el caso que nos ocupa, esta matriz contiene términos cruzados. Si elegimos de manera adecuada las variables del problema, esta matriz será diagonal, conteniendo, de esta manera, las constantes de fuerza relacionadas con cada una de las variables que hemos elegido para caracterizar nuestro problema.

El modelo de interacción binaria entre capsómeros que vamos a emplear en este trabajo viene caracterizada por un potencial de tipo *coarse-grained*. Hemos elegido este modelo dado que describe las estructuras geométricas de varios virus icosaédricos diferentes bajo criterios de mínima energía y frustración [3]. Los capsómeros que hemos elegido han sido los trímeros, que tienen forma de triángulos equiláteros. Este tipo de capsómeros se han encontrado en algunos virus de menor tamaño [4], y dado que no han sido estudiados en profundidad, hemos empleado los modelos de interacción antes mencionados para calcular las constantes de fuerza relacionados con estos cuerpos de grupo de simetría C_3 . En concreto, vamos a introducir en el potencial típico del modelo de interacción *coarse-grained* términos adicionales, dado que, como comentaremos posteriormente en este trabajo, estamos considerando sólidos rígidos con grupo de simetría C_3 . Esto va a añadir grados de libertad adicionales al problema, y, en consecuencia, aparecen nuevas constantes de fuerza relacionadas a estos grados de libertad. De esta manera, aparece un nuevo ángulo χ , que caracteriza una rotación en torno al eje Z , que surge tras romper la simetría que existe en este eje en los modelos *coarse-grained* de carácter más general.

3. Estructura de los virus icosaédricos

Abstract. *Icosahedral capsids are formed by exactly 60 subunits made of proteins. Thanks to the Caspar and Klug model, we can classify and explain the viruses icosahedral form. This model constructs the icosahedral viruses through a hexagonal lattice. Replacing 12 hexamers by 12 pentamers we can form, by Euler's theorem, a closed surface. There are counted ways to form an icosahedral capsid. The number of forms that we can use to create icosahedrons is related to the triangulation number T , defined first by Caspar and Klug. We can know the number of proteins in a capsid or the total number of capsomers in a capsid by knowing this triangulation number T .*

Como ya dijimos antes, los virus pueden tener distintas formas de cápside, de las que la estructura más común es la cápside icosaédrica. Por lo tanto, esto quiere decir que los capsómeros se van a organizar de forma que formen un icosaedro. Además, los capsómeros suelen estar formados por una única proteína. Esto puede ser debido a muchos factores. Biológicamente, la utilización de una única proteína para la formación de los capsómeros es especialmente eficiente en cuanto a la información genética, dado que una estructura regular hecha por capsómeros con la misma proteína supone una menor utilización de material genético comparado con una hipotética cápside hecha por una macroproteína [5]. Las formas de construir una estructura regular, con grupo de simetría cúbica, combinando proteínas asimétricas son limitadas, de las que la estructura icosaédrica es la mayor estructura que puede acomodar 60 subunidades asimétricas, y además, maximiza la cantidad de volumen por unidad de superficie. Pero eso no implica que debe estar hecho por exactamente 60 proteínas. Estudios mostraron que la misma simetría era compatible con cápsides hechas por distinto número de subunidades [6]. Los primeros en sentar las bases de la teoría de las cápsides icosaédricas fueron Caspar y Klug [7].

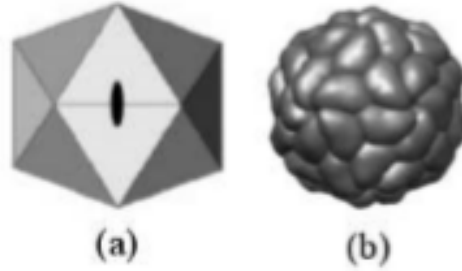


Figura 3: Cápsides icosaédricas centradas en un eje de dos dobleces. (a) Forma esquématica de una cápside. (b) Virus satélite del mosaico del tabaco. Este virus está constituido por 60 proteínas, que se agrupan en pentámeros alrededor de cada uno de los doce vértices del icosaedro [5].

3.1. Modelo de Caspar y Klug

Este modelo se basa en la construcción de cápsides icosaédricas mediante una red hexagonal plana, o de forma equivalente, una red triangular. De esta manera, al emplear esta red, hay un número concreto de formas de crear estructuras cerradas con simetría icosaédrica. Para ello, se reemplaza 12 hexámeros distribuidos de igual manera por 12 pentámeros, que son los que se requieren por el teorema de Euler para construir una superficie cerrada [8]. Para cumplir esto de diferentes formas se corresponde con los números triangulación, T . Este número es muy útil en el modelo, ya que se emplea para clasificar de manera sistemática los virus icosaédricos. El número T está definido por el vector que une dos pentámeros adyacentes en la red, de manera que:

$$\vec{C}_T = h\vec{a}_1 + k\vec{a}_2 \equiv (h, k)$$

donde (h, k) son enteros positivos que dan el número de pasos para conectar los dos pentámeros más cercanos a lo largo de las direcciones principales de la red hexagonal (ver Figura 4) La cara triangular más pequeña está definida por $(1, 0)$ o de forma equivalente por $(0, 1)$, y tiene un área $S_0 = |\vec{a}_1 \times \vec{a}_2|/2$. El número de triangulación T es el área del triángulo equilátero definido por \vec{C}_T dividido por S_0 , o de forma equivalente, el número de triángulos equiláteros de área S_0 contenidos en una cara del icosaedro resultante [5]. Operando, se obtiene la siguiente relación:

$$T = h^2 + hk + k^2$$

Dado que h y k son enteros positivos, T toma valores de la siguiente secuencia: $T = 1, 3, 4, 7, 9, 12, 13, \dots$ Además, T corresponde también al número de localizaciones quasi-equivalentes de las subunidades de la cápside [7][9].

El número total de proteínas en una cápside es:

$$N_{sub} = 60T$$

dado que para un triángulo con $T = 1$ contiene tres proteínas y el icosaedro resultante está hecho por veinte caras T . Además, cada estructura tiene $N_p = 12$ pentámeros, sumando sesenta proteínas, y las subunidades restantes se distribuyen en $N_H = 10(T - 1)$ hexámeros. Por tanto, el número total de capsómeros en la cápside es de:

$$N = 10T + 2$$

Este número T descrito por el modelo de Caspar y Klug se emplea para describir virus icosaédricos actualmente.

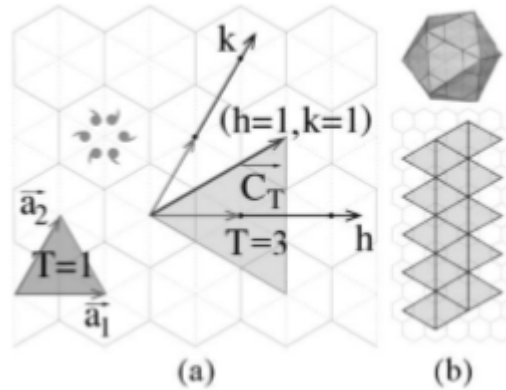


Figura 4: (a) Elementos básicos del modelo de Caspar y Klug. Esta red hexagonal se puede subdividir en pequeños triángulos que pueden ser asociados a subunidades simples. A partir de estos triángulos, se pueden construir otros triángulos de diferente tamaño, que se corresponderán a las distintas caras del icosaedro final. Además, los vértices de este triángulo estarán asociados con los vértices del icosaedro. La red está formada por los vectores \vec{a}_1 y \vec{a}_2 , relacionados con las coordenadas h y k . Estos últimos expresan los pasos a lo largo de la red. (b) Como se puede observar, si plegamos la red oscurecida sobre sí misma, formamos el icosaedro que está encima. Esta red está formada por triángulos con $T = 3$, como el que se muestra en (a) [5].

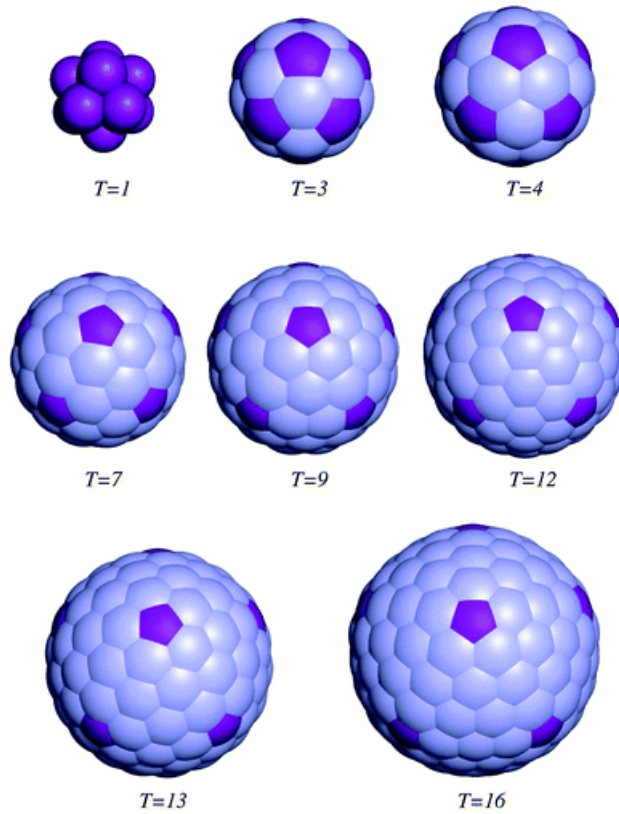


Figura 5: Ejemplos de cápsides icosaédricas correspondientes a diferentes números T . Cada una de estas estructuras con diferente número de triangulación T puede encontrarse en numerosos virus, como el picornavirus, que posee un número de triangulación $T = 3$, o el herpesvirus, con número de triangulación $T = 16$.

4. Formación de la cápside

Abstract. *In this chapter, we will explain the formation of the capsid through thermodynamics and kinetics. As we did know, virus auto-assembly their capsids. First, we considered a thermodynamical system of identical subunits that can assembly a $T=1$ icosahedral capsid. Thanks to the Mass-Action Law, we can describe the total energy needed to form a capsid minimizing the free energy of the subunit-subunit interaction. Since we know that there is only free subunits or completed capsids, we can simplify our equations. After that, we can define a critical subunit concentration which decides if the capsid forms. In the kinetics of the assembly, we will study the factors that are involved in the capsid formation, such as the subunit concentration or the interaction strength between subunits. As we should suppose first, if we increase the value of these parameters, the capsid formation process will be faster. But we have to take into account kinetics traps, that stops our capsid formation. Conclusion is capsid formation is sophisticated.*

Como ya habíamos visto anteriormente, una de las características de los virus más impresionantes es la capacidad de autoensamblaje. Para que este proceso tenga lugar, tendrá que haber subunidades proteicas cercanas entre si, además de ciertas condiciones de pH o salinidad. Después del ensamblaje, las proteínas, a veces, son modificadas en un proceso llamado maduración. Este proceso fortalece la cápside, mediante reacciones químicas y cambios conformacionales, para soportar las grandes presiones internas o protegerse de ambientes externos que podrían poner en peligro la consistencia de la cápside [10].

Este mecanismo está gobernado principalmente por la termodinámica y la cinética de ensamblaje. Empezaremos analizando la formación de cápsides vacías, dado que muchos virus suelen crear primero una cápside vacía y, más tarde, se introduce el material genético dentro de la misma.

4.1. Termodinámica del ensamblaje

Vamos a considerar un sistema termodinámico de subunidades idénticas que se pueden ensamblar en cápsides vacías con $T = 1$. Además, para simplificar la

representación, asumimos que hay una especie intermedia dominante por cada número de subunidades n . Si minimizamos la energía libre total bajo la restricción de la concentración total fija de subunidades:

$$c_T = \sum_{n=1}^N n c_n$$

nos da la Ley de Acción de Masas para la concentración de equilibrio de cada especie c_n [2]:

$$c_n v_0 = (c_1 v_0)^n e^{(-G_n^{cap}/k_B T)} \quad (1)$$

siendo v_0 el volumen de estado estándar, $k_B T$ la energía térmica y G_n^{cap} es la energía libre debida a la interacción subunidad-subunidad para el intermediario n .

Bajo la mayoría de condiciones al equilibrio, casi todas las subunidades son encontradas o en cápsides completas o libres. Esta situación resulta de cualquier modelo de ensamblado de estructuras de forma finita en los que la energía libre de interacción G_n^{cap} es mínima para una estructura ($n = N$) y la concentración total de subunidades se conserva [11]. Bajo estas condiciones la ecuación (1) puede ser simplificada despreciando todos los términos intermedios que no sean subunidades libres o cápsides completas, de manera que $c_T = c_1 + N c_N$, siendo N el número de subunidades en una cápside completa. En el límite $N \gg 1$, la fracción de subunidades en cápsides $f_c = N \cdot c_N / c_T$ viene dada por [2]:

$$f_c \approx \left(\frac{c_T}{c^*} \right)^2 \ll 1 \quad c_T \ll c^* \quad (2)$$

$$\approx 1 - \frac{c^*}{c_T} \quad c_T \gg c^* \quad (3)$$

siendo $c^* \approx v_0^{-1} \cdot e^{(G_N^{cap}/Nk_B T)}$ la concentración quasi-crítica de subunidades, por debajo de la cual no se ensamblan cápsides.

Bajo condiciones biológicas, dentro de células infectadas por el virus, la concentración de proteínas producidas por la célula tendría que exceder c^* antes de que empiece el ensamblaje de nuevas cápsides. Además, como mencionamos previamente, la Ley de Acción de Masas es consecuencia de la minimización de la libre energía de Gibbs, de forma que requiere proteínas necesarias para formar la cápside en una solución con el mismo potencial químico que las proteínas incorporadas en una cápside ya formada [10]. Sin embargo, cuando la concentración total de proteínas se reduce de nuevo por debajo de c^* después de completar el ensamblaje, entonces las cápsides deberían desmontar de forma espontánea de

acuerdo a la Ley de Acción de Masas. En realidad, esto no es lo que pasa, o pasa después de un periodo de tiempo muy largo, o cuando varía sustancialmente el pH, la salinidad u otras condiciones de la solución [12]. Este 'exceso' de estabilidad termodinámica en las cápsides ya formadas, cuando lo comparamos con el equilibrio convencional de autoensamblaje, es, desde un punto de vista biológico, una característica primordial de supervivencia, dado que los virus deben permanecer intactos en ambientes 'hostiles' donde no hay proteínas para ensamblar virus, como el torrente sanguíneo o los ácidos estomacales [10].

4.2. Cinética del ensamblaje

La cinética del ensamblaje de cápsides, sea medida experimentalmente o calculada por modelos teóricos o computacionales, tiende a formar una curva sigmoide [2]. Comienza con una fase de latencia en la que se forman las fases previas de la cápside. Seguidamente, este proceso de latencia es seguido por una producción rápida de cápsides, que más tarde se aproximará asintóticamente al equilibrio. En esta última fase, el ensamblado de cápsides se ve reducido mientras que las barreras de nucleación incrementan debido al agotamiento de las subunidades libres [2]. Si incrementamos la concentración de subunidades c_T o la intensidad de las interacciones internas entre subunidades g_b , mediante la disminución del pH o incrementando la concentración de sal en la solución, inicialmente lleva a una formación de las cápsides más rápida. Sin embargo, mientras la termodinámica (Ecuaciones 2 y 3) indica que el rendimiento de las cápsides bien formadas incrementa monótonamente con g_b y c_T , el rendimiento en tiempos largos no incrementa de esta manera debido a las trampas cinéticas, es decir, a las distintas condiciones que se pueden dar a la hora de la formación de las cápsides que impiden o malforman la cápside [2].

Los estudios experimentales y los modelos muestran que hay una compensación entre la especificidad de interacción y la accesibilidad cinética: interacciones más específicas incrementan la selectividad del montaje para la estructura objetivo, pero disminuye el ratio de ensamblaje debido a la disminución cinética de las secciones cruzadas [13][14]. Los resultados del enfrentamiento entre la termodinámica y la cinética se resumen en la figura 6, donde se describe el resultado predominante a tiempos largos como función de la concentración de subunidades y la intensidad de la interacción. Para una especificidad de interacción dada, el diagrama de fases cinético puede ser clasificado en cinco partes [2].

- En esta primera fase, el montaje de cápsides es no favorable en el equilibrio, debido a las interacciones débiles o concentraciones bajas de subunidades como en el caso de $c_T < c^*$ (Ecuación 2).
- En esta parte, mientras las interacciones o la concentración de subunidades incrementan a $c_T \geq c^*$, el montaje de cápsides es favorable, pero no ocurre experimentalmente para escalas de tiempo relevantes debido a las grandes barreras de nucleación.
- En este caso, incrementando notablemente las interacciones o la concentración de subunidades, nos lleva a barreras de nucleación moderadas y grandes rendimientos en la formación de cápsides en escalas de tiempo relevantes (que puede llevar de segundos a horas para formar cápsides vacías [2]). Finalmente, interacciones por encima de lo óptimo llevan a empeorar los rendimientos debido a dos trampas cinéticas:
 - ◇ Cuando la nucleación es rápida comparado con el crecimiento de la cápside, muchas cápsides nuclean de forma temprana y las subunidades libres o las pequeñas estructuras formadas se agotan antes de que un número significativo de cápsides finalice el ensamblaje [15][16]. Esto suele ocurrir cuando las escalas de tiempo requeridas para que las cápsides completen la fase de crecimiento exceden las escalas de tiempo de nucleación [17][18].
 - ◇ Bajo interacciones lo suficientemente fuertes, subunidades con orientaciones incorrectas son atrapadas en estructuras formadas en crecimiento por adiciones posteriores de subunidades, llevando o, a cápsides defectuosas ya cerradas a falta de simetría icosaédrica, o, a estructuras espirales abiertas tanto en simulaciones como en experimentos [15][16][19].

La presencia de estas dos formas de trampas cinéticas explican algunas observaciones que indican que se requieren interacciones débiles para el montaje productivo de cápsides virales [20][15].

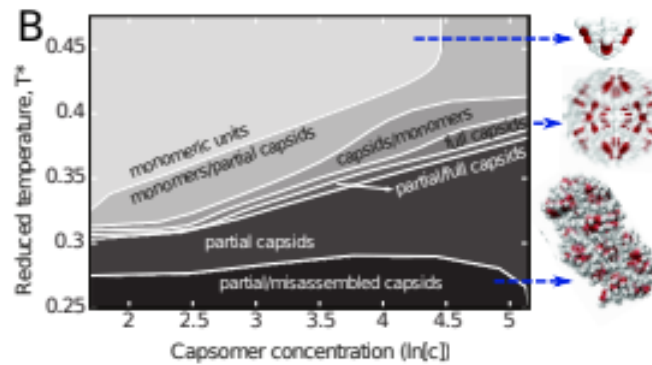


Figura 6: Diferentes productos derivados del ensamblaje de las cápsides con tiempos grandes para una cápside icosaédrica de 20 subunidades como función de la temperatura (inversa de la intensidad de interacción) y concentración de subunidades. A la derecha de la imagen se muestran estructuras representativas de las diferentes etapas de la formación de cápsides virales [2]. Esta imagen ha sido tomada de [2].

5. Propiedades físicas de los virus

Abstract. *In this chapter, we will study the electrostatics and mechanical properties of viruses. First, we can consider the capsomers positively charged, but if only this electrostatic interaction were working, we could not have auto-assembly, since the capsomers would repel. But as we know there are other forces interacting, such as Van der Waals interaction and hydrophobic interaction, being the last the dominant interaction. We can consider the capsid a positively charged sphere with a superficial charge density s with a radius r . Using Poisson-Boltzmann's equation, we can describe physically this electrostatic problem. Doing some approximations, we can obtain the self-energy F of the capsid. Mechanical properties of the capsid is another thing we can study. We can apply a external force via nano-indentation with AFM. If we consider some parameters to describe physically the capsid, such as the osmotic pressure generated by the genome or the radial inwards displacement, we can obtain an expression for the deformation free energy, but first, we must use thin-shell theory to interpret measured force-deformation curve, which is related to the deformation free energy.*

Las cápsides virales son objetos que llevan siendo estudiados físicamente a lo largo de la historia hasta la actualidad. Esto es debido a diferentes características físicas que lo describen, como la capacidad electrostática de los capsómeros para ensamblarse, la distribución de carga en la cápside, así como las propiedades mecánicas y elásticas de un virus midiendo la resistencia que ofrece al aplicarle una fuerza externa.

5.1. Electroestática

Como muchas moléculas libres, las proteínas que forman los capsómeros poseen carga eléctrica positiva. De este modo, la representación más simple de la distribución de carga de una proteína puede ser una distribución homogénea de carga positiva. Sin embargo, si esto solo fuera así, el autoensamblaje nunca tendría lugar, dado que las proteínas se repelerían entre sí, pero está demostrado empíricamente que los virus se autoensamblan, por lo que existirán otras fuerzas de interacción entre capsómeros que permitan el autoensamblaje. Estas fuerzas son

una combinación entre la interacción de Van der Waals y la hidrofóbica, siendo esta última la interacción dominante [1].

Una cápside viral puede representarse como una esfera uniforme, cargada positivamente con una densidad superficial de carga s , permeable, infinitamente fina y de radio R . Empleando la ecuación de Poisson-Boltzmann (PB), una ecuación diferencial no lineal, se puede obtener la energía propia electrostática, dado que incorpora información detallada sobre la forma biomolecular y la distribución de carga [1]. Esta ecuación puede derivarse minimizando la energía libre apropiada:

$$F_{PB}[\Phi(r), \nabla\Phi(r), c^i(r)] = \int f_{PB}(\Phi(r), \nabla\Phi(r), c^i(r)) d^3r \quad (4)$$

Si consideramos los iones como gas ideal, los potenciales electrostáticos en la solución pequeños y si el sistema es simétrico, las ecuaciones de Poisson-Boltzmann pueden ser linealizadas, lo que nos lleva a la ecuación de Debye-Huckel (DH) para el potencial:

$$-\nabla^2\Phi(r) = \frac{\beta}{\epsilon\epsilon_0} \left(\sum_{i=\pm} e_i^2 c_0^i \right) \Phi(r) + \frac{e_0\rho_p(r)}{\epsilon\epsilon_0} + \dots \quad (5)$$

El parámetro $e_0\rho_p(r)$ es la densidad de carga de la cápside, c_0^i es la concentración de bultos de iones de sal, $\epsilon\epsilon_0$ es la permitividad del agua, $\beta^{-1} = k_B T$, siendo T la temperatura y k_B la constante de Boltzmann, e_i es la carga de los iones, de forma que $e_i = \pm e_0$, y e_0 es la carga del electrón [1].

La ecuación de Debye-Huckel puede ser resuelta de varias maneras, dado que es una ecuación lineal. Se puede obtener la energía propia de una esfera cargada uniformemente sumando el par de interacciones Debye-Huckel sobre la superficie de la esfera, de forma que:

$$F_{DH} = \frac{1}{2} \frac{\sigma^2}{4\pi\epsilon_0\epsilon} \int dS_1 \int dS_2 \frac{e^{-\kappa|\vec{r}_1 - \vec{r}_2|}}{|\vec{r}_1 - \vec{r}_2|} \quad (6)$$

donde dS_1 y dS_2 son elementos infinitesimales de la superficie de la esfera centrado en torno a r_1 y r_2 . El factor $\frac{1}{2}$ es para eliminar uno de los términos que resulta de contar dos veces el par de interacciones. Por otro lado, se debe tener en cuenta que el radio de los virus son, algunos, del orden de 20 nm, mientras que la longitud de detección de Debye-Huckel, en condiciones fisiológicas, es $\kappa^{-1} \approx 1$ nm. Si $\kappa R \gg 1$, eso significa que el rango de integración puede cortarse en la escala de

κ^{-1} . Cuando $\kappa R \gg 1$, en vez de definir una esfera, la integral define un disco, haciendo los dos integrandos independientes [1]. Desarrollando la expresión, queda la energía propia de la cápside:

$$\lim_{\kappa R \gg 1} F_{DH} = \frac{\pi \sigma^2 R^2}{\epsilon_0 \epsilon \kappa} \quad (7)$$

Por lo tanto, la energía propia F es la energía necesaria para traer cargas infinitesimales desde separaciones infinitas en la solución hacia la cápside. La cantidad puede considerarse como la contribución electrostática a la libre energía de ensamblado. Para una cápside de alrededor de $\sigma \approx 1 \text{ e}_0/\text{nm}^2$, y $c_0 = 100 \text{ mM}$, la libre energía sería $F_{DH} \approx 10^4 \text{ k}_B T$. Una cápside de alrededor de 20 nm es típicamente construida por 180 subunidades proteicas, lo cual hace que la longitud de los contactos entre proteínas sea alrededor de 3000 nm . Si la cápside tiene un grosor de 2 nm , el área total expuesta a los contactos entre proteínas es alrededor de 6000 nm^2 . Teniendo en cuenta que la energía de las interacciones atractivas de las proteínas por unidad de área de la superficie de la proteína expuesta es del orden de 10 mJ/m^2 , la interacción atractiva hidrofóbica es de $F_{DH} \approx 10^4 \text{ k}_B T$ [1]. Como se puede observar, las energías de interacción calculadas anteriormente coinciden en magnitud, lo cual puede explicarse por el hecho de que los virus, a parte de necesitar ensamblarse, también necesitan desensamblarse para liberar el material genético a la célula huésped.

El modelo descrito anteriormente se puede refinar. Para ello, si se examina todo el rango de κR , se considera que el potencial electrostático $\beta \epsilon_0 \phi$ no es tan pequeño como para despreciarlo, añadimos un grosor finito δ a la cápside, y si se añaden las condiciones $\kappa R \gg 1$, $\epsilon \gg \epsilon_p$, $\delta \ll R$ y $\kappa \delta \approx 1$, obtenemos la siguiente ecuación para la energía propia F [1]:

$$F_{DH} = \frac{2\pi(\sigma_1^2 + \sigma_2^2)R^2}{\epsilon_0 \epsilon \kappa} \quad (8)$$

donde σ_1 es la densidad de carga superficial de la cara interna de la cápside y σ_2 es la densidad de carga superficial de la cara externa de la cápside.

5.2. Propiedades mecánicas

Es interesante comprender la resistencia que tienen los virus, o cápsides vacías después del autoensamblado, cuando se les aplica una fuerza externa, dado

que necesitan ser lo suficientemente resistentes para ser estables para proteger el genoma interno en el ambiente extracelular, y lo suficientemente inestable como para liberar las moléculas genéticas en las células huésped para más tarde reproducirse. Numerosos experimentos se han llevado a cabo en virus para determinar sus propiedades mecánicas, y el más común es el microscopio de fuerza atómica, que prueba los virus y localiza deformaciones posteriormente [21].

La relación entre la fuerza aplicada y el cambio que resulta en el diámetro de la cápside es llamada la curva de la fuerza de deformación. Por otro lado, dependiendo de si la cápside vuelve o no a su forma original después de aplicar una fuerza, se tratará de una deformación reversible o irreversible, respectivamente. La fuerza es medida mediante sondeos de nanoindentación (ver Figura 7) que resulta del hecho que la prueba aleja la cápside del estado de mínima energía libre. Sin embargo, para interpretar las curvas de fuerzas de deformación medidas, es necesario hacer una analogía con la teoría de envolventes delgadas. Esta teoría es utilizada para predecir los efectos de las fuerzas externas en paredes delgadas o estructuras huecas macroscópicas, pero puede aplicarse de igual manera a un virus, si la cápside es modelada de forma fina, esférica y con un radio R [10]. Si tenemos en cuenta el genoma del virus, debe saberse que existe en consecuencia una presión osmótica Π que debe incluirse en los cálculos, dado que puede alcanzar una presión de 50 atmósferas [22][23]. Esta presión interna va a estar relacionada con la tensión no específica τ a lo largo de la cápside de acuerdo con la ley de Laplace $\Pi = \frac{2\tau}{R}$, lo cual incrementa la constante elástica del virus [10].

Si la sonda de fuerza deja un perfil de indentación $\zeta(r)$, definido como el desplazamiento radial interno de la superficie de la esfera expresado en términos de un sistema de coordenadas bidimensional que cubre la cápside, pasa a ser muy pequeña, la energía libre de deformación de la teoría de envolventes delgadas ΔF es función de $\zeta(r)$ en la forma de una integral a lo largo de la superficie de la cápside [10]:

$$\Delta F = \int dS \left(\frac{1}{2} \kappa (\Delta \zeta)^2 + \frac{1}{2} \tau (\nabla \zeta)^2 + \frac{1}{2} Y \left(\frac{2\zeta}{R} \right)^2 \right) \quad (9)$$

donde el primer término describe el coste energético de flexión de la indentación, el segundo término representa el trabajo por la sonda contra la presión osmótica que ejerce el genoma Π . El tercer término mide la extensión de la capa inducida por la fuerza con el módulo de Young bidimensional Y de la capa [10]. Para obtener la curva de la fuerza de deformación, se hace la derivada funcional

de la energía libre de deformación, resultando en:

$$\frac{\delta \Delta F}{\delta \zeta(\vec{r})} = f(\vec{r}) \quad (10)$$

donde $f(\vec{r})$ es la fuerza radial por unidad de área ejercida por la sonda. La ecuación 10 puede resolverse analíticamente en el caso de una fuerza puntual, con $f(\vec{r}) = F \delta(\vec{r})$, que crea una curva de la fuerza de deformación lineal y un hoyo con un radio de orden $\sqrt{Rl_b}$, siendo $l_b = \sqrt{\kappa Y}$ [10]. Para fuerzas aplicadas pequeñas, la cápside se comporta como un oscilador armónico y la curva de la fuerza de deformación medida es, como suponíamos, pequeña en muchos casos [10].

Las indentaciones mayores son más complicadas. Para calcular la curva de las fuerzas de deformación de la teoría de envolventes delgadas en el régimen no lineal, requiere la solución de dos ecuaciones diferenciales no lineales llamadas ecuaciones de Föppl - von Kármán (FvK). En vez de resolver las ecuaciones de FvK analíticamente o numéricamente, es más práctico minimizar numéricamente la energía elástica usando directamente modelado de elementos finitos (*finite-element modelling*, FEM). Las conclusiones recogidas por el modelado de elementos finitos muestran que la respuesta elástica de las cápsides icosaédricas no uniformes pueden diferir de las cápsides esféricas uniformes [10].

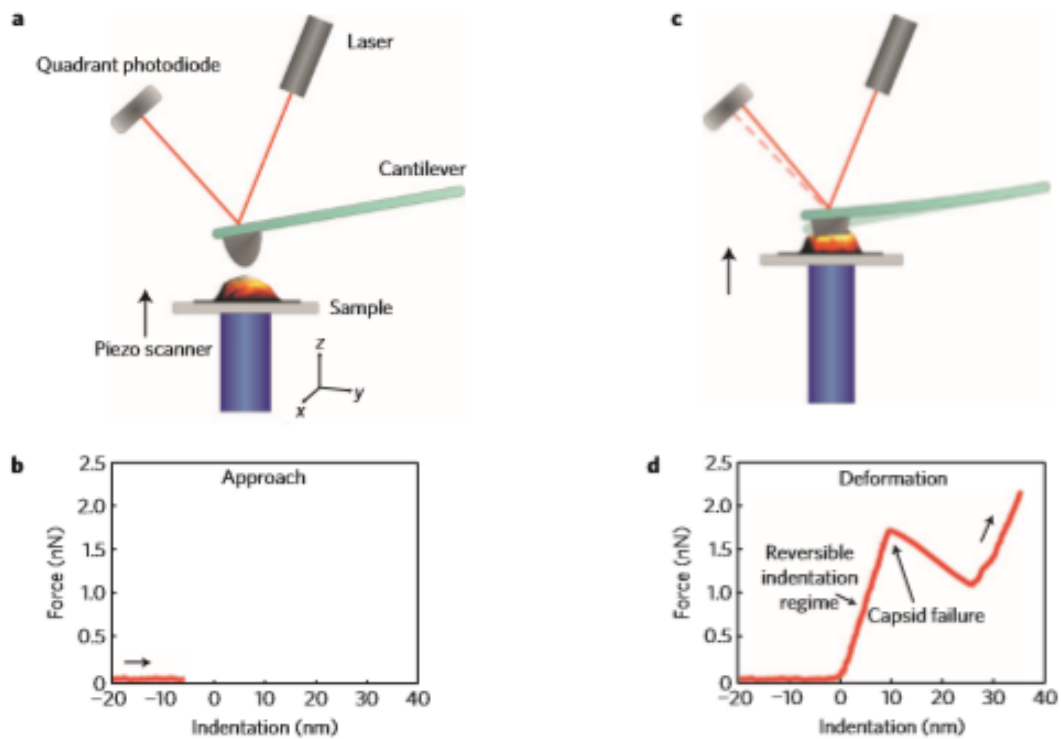


Figura 7: Diagrama esquemático de un experimento de nanoindentación en un virus. El radio de curvatura del microscopio es de 2 – 20 nm, un tamaño menor o comparable a un virus de pequeño tamaño. En **b** se puede observar como la fuerza aplicada es nula, dado que aún no se está aplicando una fuerza al virus (**a**). El microscopio ya está indentando el virus (**c**), mientras en **d** podemos ver como la fuerza se aplica al virus en diferentes etapas del proceso: se muestra una deformación lineal con pendiente positiva, ya que se aumenta la fuerza aplicada; le sigue una deformación con pendiente negativa, que suele ser irreversible para la cápside, dado que ésta se rompe y libera parte de la fuerza aplicada [10].

6. Modelos de interacción

Abstract. *We will study the two models that describes the interaction potential between capsomers. These models are Coarse-grained and All-atom. All atom models consider every atom of the system, so this model will use quantum mechanics to explain the interactions between the capsomers. Since a virus, or a capsid, is a very complex system, because of the amount of atoms that interact between them, and the computational times that it requires to solve these systems are quite long, this model is very complex, but its quite accurate. Coarse-grained models group all this atoms into a system of a few particles, such as capsomers. This supposition reduces the amount of degrees of freedom of our problem, and so, we can consider only the degrees of freedom that 'really' characterizes a system of particles. This model uses approximations, but it is still accurate, and reduces the amount of time needed to characterize a physical system.*

En secciones anteriores se ha descrito el funcionamiento del autoensamblaje de cápsides virales, así como las fuerzas electroestáticas que tienen lugar entre los propios capsómeros libres o la interacción que tiene lugar entre ellos. Esta interacción puede ser modelada en potenciales de interacción. Existen dos tipos de modelos para describir el potencial de interacción entre capsómeros o entre las unidades fundamentales del virus: *Coarse-grained* y *All-atom*.

6.1. Modelos *All-atom*

La teoría más precisa para describir procesos de escala atómica es la mecánica cuántica. Resolviendo, por ejemplo, la ecuación de Schrödinger para un sistema de partículas, podemos obtener la evolución temporal de las distribuciones de probabilidad para las posiciones de electrones y núcleos. Sin embargo, cuantas más partículas se le añadan al sistema, más complejo se vuelve. El modelo *all-atom* se basa en el modelado de moléculas de diferente escala mediante la representación de cada átomo que conforma el sistema global. Ofrece gran precisión en el modelado de sistemas de partículas que interactúan entre sí, dado que trabaja con sistemas cuánticos, los cuales se rigen por la mecánica cuántica. Pero, por otro lado, la complejidad del sistema va a requerir gran cantidad de tiempo computacional,

dado que una biomolécula puede contener una inmensa cantidad de partículas, las cuales interactúan todas entre sí. Usando ordenadores altamente especializados, las simulaciones basadas en mecánica molecular han sido extendidas a escalas de tiempo del milisegundo para sistemas pequeños como la ubiquitina, con 10 mil átomos en el modelo [24]. Uno de los modelizados más grandes que se han hecho con este modelo es el de la cápside del virus VIH-1, el cual contiene 64 millones de átomos en el sistema modelizado. Para obtener trayectorias dinámicas de 0,1 *vs* se necesitó dos semanas de tiempo computacional en 128 mil núcleos [24]. Aunque este modelo es más preciso que el *coarse-grained*, supone un gran desafío estudiar sistemas tan complejos. De cualquier forma, con los avances en computación cuántica, estas simulaciones de modelizado podrán realizarse con un rendimiento mucho mayor, ya que reducirá en gran cantidad el tiempo computacional, y, por tanto, el modelo *all-atom* tendrá cada vez más impacto en el campo de la investigación.

6.2. Modelos *Coarse-grained*

El modelo *Coarse-grained* intenta generalizar el comportamiento de un grupo o colectivo de partículas o cuerpos, ya que despreja el comportamiento individual de los átomos, para así reducir la complejidad del sistema de partículas. Así, este modelo va a tener en cuenta el comportamiento de un sistema a gran escala, sin importar las fluctuaciones individuales de cada átomo. Por lo tanto, podemos aumentar el alcance de nuestras simulaciones, así como reducir los tiempos computacionales. Una forma de hacer esto es agrupando gran cantidad de átomos y representándolos como partículas simples. Sin embargo, este modelo no se emplea únicamente para acelerar procesos de cálculo computacional. Otra razón que lleva a emplear este modelo es para ser capaces de tener los grados de libertad que son relevantes para nuestro problema, y también para ajustar las interacciones a esa escala, permitiéndonos estudiar los efectos de eliminar grados de libertad o cambiar las interacciones a diferentes escalas [24]. Por otro lado, este modelo nos ofrece un gran rango de escalas, pudiendo elegir el más adecuado a nuestro problema, pero intenta evitar entrar en escalas de nivel cuántico, dado que a este nivel aparecen nuevas cuestiones a tener en cuenta que volverían nuestro problema realmente complejo.

Este modelo se usa para el modelado molecular de biomoléculas a distintos niveles de granularidad, así como para el modelado computacional de moléculas específicas: proteínas, ácidos nucleicos, membranas lipídicas, carbohidratos o agua.

Las primeras aproximaciones de este modelo proviene del trabajo de Michael Levitt y Ariel Warshel en la década de 1970 [25]. Estos investigadores fueron los primeros en llevar el plegamiento de una proteína a una computadora, por medio de una combinación de procedimientos de cálculo por minimización de energía y termalización, utilizando una representación estructuralmente simplificada de una proteína. Su modelo de proteína representaba los muchos átomos contenidos en cada aminoácido de la cadena polipeptídica con tan sólo dos centros, uno para la parte del esqueleto proteico y otro para la cadena lateral. Este tipo de modelo en el que la estructura proteica es sólo descrita a un nivel de baja resolución, fue más tarde refinado y extensivamente empleado en simulaciones de plegamiento que han proporcionado numerosos descubrimientos a la comunidad científica.

Aunque este modelo es menos preciso que el *all-atom*, es uno de los modelos que más se ha utilizado para estudiar y describir la interacción entre capsómeros, debido a que se ajusta fielmente a los resultados experimentales. Por este motivo, se ha decidido emplear el modelo *coarse-grained* en este trabajo para describir la interacción entre dos trímeros y calcular las constantes de fuerza relacionadas.

7. Aplicación práctica

Abstract. *In this chapter, we will use a interaction potential to describe our problem of two trimers interaction. The model selected is Coarse-grained. First, we consider a general Lennard-Jones potential, that only depends on the radius. Afterwards, we will introduce the variables that characterize the trimer orientation and the equilibrium conditions that fixes the privileged orientation of the trimers in order to form a capsid. Then, we will introduce parameters in the interaction potential that characterize the orientation of the trimer, and we will introduce as well the Euler angles related with the variables that describes the movement of the trimer related with the degrees of freedom of a trimer. After that, we will calculate the second derivative matrix of the interaction potential respect the variables that describes our problem in order to calculate the force constants. Finally, we apply the equilibrium conditions to the matrix to obtain the force constants.*

Antes de comenzar el cálculo de las constantes de fuerza, vamos a explicar el potencial de interacción que hemos utilizado.

7.1. Potencial de interacción

Como ya comentamos en el apartado 2.1 de objetivos, nuestro modelo de potencial de interacción será *coarse-grained*.

Vamos a desarrollar la forma de la interacción entre dos cuerpos anisotrópica escribiéndolo una expansión truncada en la base completa de los términos angulares de las expansiones multipolares estándar, ya que incorpora las características geométricas más importantes de la interacción binaria entre capsómeros en cápsides de virus icosaédricos [3]. Nuestro modelo utiliza el vector posición \mathbf{r}_i que nos da las coordenadas al centro geométrico del capsómero i en un espacio tridimensional. Por otro lado, introducimos el vector \mathbf{v}_i que da la orientación del capsómero en la dirección perpendicular del plano principal del capsómero. Dado que la interacción entre dos capsómeros viene caracterizada mediante la orientación angular de éstos, se va a emplear una expansión multipolar mediante una serie matemática. Los términos en esta expansión añaden progresivamente características angulares más finas, y se les llama según el orden del momento de la expansión multipolar.

Por ejemplo, el término isotrópico más simple es el monopolo-monopolo, y siguiendo progresivamente la complejidad tenemos el monopolo-dipolo, el dipolo-dipolo, el monopolo-cuadrupolo, o el dipolo-cuadrupolo. La dependencia orientacional de la interacción entre dos capsómeros (i,j) con el vector de posición relativo $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$ junto con $\mathbf{n}_{ij} = \mathbf{r}_{ij}/r_{ij}$ utilizará los términos multipolares más simples en modelos más generales [3]. En nuestro caso, se utilizará un desarrollo multipolar más complejo, dado que los capsómeros que hemos escogido han sido los trímeros, los cuales poseen un grupo de simetría C_3 . De esta manera, al introducir esta geometría, se rompe con la simetría que existía en rotaciones en torno al eje z , ya que en modelos más generales, los capsómeros tienen forma de discos. Así, aparece un nuevo grado de libertad relacionado con esta ruptura de la simetría. Por lo tanto, hará falta una variable más en el problema para caracterizarlo completamente. Esta variable es el ángulo χ , que describe la rotación del trímero en torno al eje z . Por lo tanto, vamos a tener 3 variables angulares para describir el movimiento de los capsómeros en nuestro problema: ϕ_i, θ_i y χ_i . Estos ángulos se pueden identificar con los ángulos de Euler. Por otro lado, vamos a añadir varios términos necesarios para que nuestro resultado cumpla con las condiciones de equilibrio, ya que los capsómeros, a la hora de unirse para formar la cápside, tienen una orientación privilegiada.

Entre las distintas formas funcionales de modelo de potencial de interacción para una primera aproximación, hemos elegido el potencial generalizado de Lennard-Jones [3], que depende solamente de la distancia entre capsómeros, pero haciendo una pequeña modificación para ponerlo en función de la distancia de equilibrio:

$$V(r) = \varepsilon \left\{ \left(\frac{r_e}{r} \right)^{2m} - 2 \left(\frac{r_e}{r} \right)^m \right\} \quad (11)$$

donde r es la distancia entre capsómeros, ε es la profundidad del pozo de potencial, m es un número entero que fija la anchura del pozo de potencial, y r_e es la distancia de equilibrio entre dos capsómeros, es decir, la distancia que hay entre ellos para que sea posible la formación de la cápside viral.

Las expansiones multipolares habituales se realizan a largas distancias y, por lo tanto, explican los términos del potencial con un decaimiento más lento a medida que r tiende a infinito [3]. Por ello, una elección conveniente de un parámetro en la ecuación 11 para incluir los términos multipolares correspondientes a nuestro problema, es el parámetro de energía ε en el segundo término de carácter atractivo y de largo alcance de la misma ecuación 11 [3].

Por lo tanto, la expresión matemática del potencial de interacción trímero-

trímero, teniendo en cuenta lo dicho anteriormente, será:

$$\begin{aligned}
 V_{tt} = \varepsilon \left\{ \left(\frac{r_e}{r} \right)^{2m} - 2 \left(\frac{r_e}{r} \right)^m \cdot \left\{ 1 + \alpha \cdot [(\mathbf{v}_{iz}\mathbf{v}_{jz}) - 2(\mathbf{v}_{iz}\mathbf{n}_{ij})(\mathbf{v}_{jz}\mathbf{n}_{ij}) - 1] \right. \right. \\
 - \frac{\gamma}{2} \cdot [(\mathbf{v}_{iz}\mathbf{n}_{ij} + \cos \theta_e)^2 + (\mathbf{v}_{jz}\mathbf{n}_{ij} - \cos \theta_e)^2] + \zeta \cdot [(\mathbf{v}_{iz}\mathbf{n}_{ij} + \cos \theta_e)(\mathbf{v}_{jz}\mathbf{n}_{ij} - \cos \theta_e)] \\
 + \eta \cdot \left[Re \left\{ (\mathbf{v}_{ix}\mathbf{n}_{ij} + i\mathbf{v}_{iy}\mathbf{n}_{ij})^3 \right\} - \left(1 - (\mathbf{v}_{iz}\mathbf{n}_{ij})^2 \right)^{3/2} \right] \\
 - \eta \cdot \left[Re \left\{ (\mathbf{v}_{jx}\mathbf{n}_{ij} + i\mathbf{v}_{jy}\mathbf{n}_{ij})^3 \right\} + \left(1 - (\mathbf{v}_{jz}\mathbf{n}_{ij})^2 \right)^{3/2} \right] \\
 \left. \left. - \delta \cdot \left[Im \left\{ (\mathbf{v}_{ix}\mathbf{n}_{ij} + i\mathbf{v}_{iy}\mathbf{n}_{ij})^3 \right\} \cdot Im \left\{ (\mathbf{v}_{jx}\mathbf{n}_{ij} + i\mathbf{v}_{jy}\mathbf{n}_{ij})^3 \right\} \right] \right\} \right\} \quad (12)
 \end{aligned}$$

Los parámetros introducidos en los distintos desarrollos multipolares del potencial: α , γ , $\cos \theta_e$, ζ , η y δ , corresponden a parámetros que caracterizan el potencial de interacción. En concreto, los parámetros α y γ reflejan los ángulos θ y ϕ que caracterizan las rotaciones en torno al eje X y en torno al eje Y , respectivamente. El parámetro ζ tiene en cuenta los parámetros cruzados entre los ángulos θ y ϕ , que hace que no se cuente doble las rotaciones de los dos cuerpos que realizan de forma coordinada en el mismo sentido, es decir, como si consideráramos el sistema de dos cuerpos como uno y lo hiciéramos rotar sobre los ángulos mencionados. Los parámetros η y δ se corresponden con términos que añaden la rotación del ángulo χ a la interacción binaria entre trímeros, dado que, como comentamos anteriormente, hay un grado de libertad más que debe ser tenido en cuenta en el potencial de interacción.

Los vectores \mathbf{v}_x , \mathbf{v}_y y \mathbf{v}_z se pueden identificar con las columnas de la matriz de rotación en función de los ángulos de Euler, siendo las componentes de \mathbf{v}_x los elementos de la primera columna de la matriz, las componentes de \mathbf{v}_y los elementos de la segunda columna y siendo las componentes de \mathbf{v}_z los elementos de la tercera columna de la matriz. La matriz en cuestión:

$$\begin{pmatrix} \cos \phi \cos \theta \cos \chi - \sin \phi \sin \chi & -\cos \phi \cos \theta \sin \chi - \sin \phi \cos \chi & \cos \phi \sin \theta \\ \sin \phi \cos \theta \cos \chi + \cos \phi \sin \chi & -\sin \phi \cos \theta \sin \chi + \cos \phi \cos \chi & \sin \phi \sin \theta \\ -\sin \theta \cos \chi & \sin \theta \sin \chi & \cos \theta \end{pmatrix}$$

Ahora, teniendo los vectores \mathbf{v}_x , \mathbf{v}_y y \mathbf{v}_z en función de los ángulos de Euler,

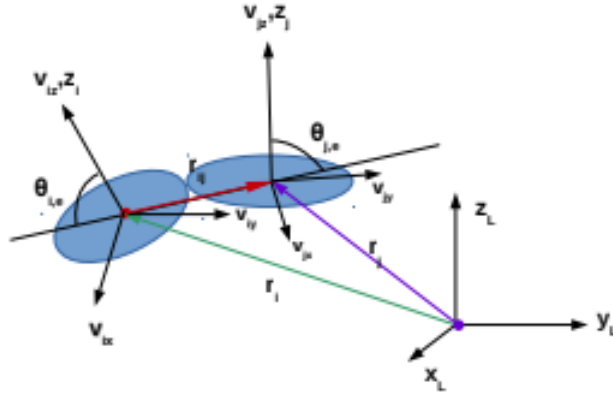


Figura 8: Representación simplificada de la interacción entre dos capsómeros con forma de disco. Como se puede observar, hay tres sistemas de referencia: el de laboratorio, l , el del capsómero i y el del capsómero j . El ángulo θ_e , como ya comentamos en el texto, es el ángulo de preferencia de los capsómeros a la hora de unirse para formar la cápside. Imagen sacada de [26].

vamos a escribir nuestro potencial de interacción en función de los ángulos de Euler:

$$\begin{aligned}
 V_{it}(\theta, \phi, \chi) = \varepsilon \cdot \left\{ \left(\frac{r_e}{r} \right)^{2m} - 2 \left(\frac{r_e}{r} \right)^m \cdot \left\{ 1 + \alpha \cdot [\sin \theta_i \sin \theta_j \cos(\phi_i - \phi_j) - \cos \theta_i \cos \theta_j - 1] \right. \right. \\
 \left. \left. - \frac{\gamma}{2} [\cos^2 \theta_i + \cos^2 \theta_j + 2 \cos^2 \theta_e + 2 \cos \theta_e (\cos \theta_i - \cos \theta_j)] \right. \right. \\
 \left. \left. + \zeta [\cos \theta_i \cos \theta_j - \cos \theta_i \cos \theta_e + \cos \theta_j \cos \theta_e - \cos^2 \theta_e] \right. \right. \\
 \left. \left. + \eta [-\sin^3 \theta_i \cos 3\chi_i - \sin^3 \theta_i + \sin^3 \theta_j \cos 3\chi_j - \sin^3 \theta_j] - \delta \cdot [\sin^3 \theta_i \sin^3 \theta_j \sin 3\chi_i \sin 3\chi_j] \right\} \right\}
 \end{aligned}
 \tag{13}$$

7.2. Cálculo de las constantes de fuerza

Como habíamos mencionado en los Objetivos del trabajo (ver 2.1), las constantes de fuerza son constantes de proporcionalidad entre una fuerza aplicada y la distancia que recorre en la dirección de la fuerza aplicada tras aplicarle dicha

fuerza. Para calcular las constantes de fuerza de nuestro potencial de interacción binaria, calculamos la matriz de derivadas segundas del potencial de interacción.

Una vez calculada la matriz, se deben aplicar las condiciones de equilibrio de nuestro problema, dado que, como explicamos en el apartado anterior, los trímeros se agruparán de acuerdo a unas orientaciones privilegiadas para formar la cápside viral correspondiente. Estas condiciones de equilibrio están relacionadas con los ángulos que describen la orientación de dos trímeros.

En primer lugar, si el contacto entre los dos capsómeros ocurre en el límite de los planos del capsómero, su geometría de equilibrio tendrá un eje de simetría doble que biseca el ángulo diédrico entre los planos, y, por este motivo, se debe tener en el equilibrio que $\phi_i - \phi_j = 0$ y que $\theta_i + \theta_j = \pi$ [3]. De este modo, si consideramos el trímero j como capsómero privilegiado, en el sentido de que se encuentra en equilibrio, podemos considerar las siguientes condiciones de equilibrio:

- La primera condición es que el ángulo θ_j se corresponda con el valor absoluto del ángulo de equilibrio θ_e , y, por tanto, el ángulo θ_i se corresponderá con el valor absoluto del ángulo $\pi - \theta_e$.
- La segunda condición es que la distancia entre capsómeros sea la distancia de equilibrio r_e .
- Dado que la diferencia entre los ángulos ϕ debe ser 0, esto significa que deben tener el mismo valor, por lo que pueden tener cualquier valor mientras sea el mismo para ambos. Vamos a darle a ambos el valor $\phi_i = \phi_j = 0$.
- Por último, la última condición, dado que el trímero j se encuentra en equilibrio, no habría que realizarle ninguna rotación en torno al eje z , por lo que $\chi_j = 0$. Ahora, para que el trímero i 'encaje' con el trímero j , dado que son triángulos equiláteros que parten de la misma posición, vamos a hacer una rotación de π en torno al eje z (otras opciones son posibles, pero ésta es la opción que más facilita los cálculos), por lo tanto, tenemos que $\chi_j = \pi$.

Tras hacer los correspondientes cálculos y diagonalizar la matriz seleccionando las variables adecuadas para que ésta sea diagonal (ver Apéndice A), se obtienen las siguientes constantes de fuerza:

Variable	Constante de fuerza	Tipo
r	$2\epsilon m^2 / r_e^2$	<i>Stretching</i>
$\phi^- = \phi_i - \phi_j$	$2\epsilon\alpha \sin^2 \theta_e$	<i>Torsión α</i>
$\theta^+ = \theta_i + \theta_j$	$\epsilon(2\alpha + (\gamma - \zeta) \sin^2 \theta_e)$	<i>Bending simétrico</i>
$\theta^- = \theta_i - \theta_j$	$\epsilon(\gamma + \zeta) \sin^2 \theta_e$	<i>Bending antisimétrico</i>
$\chi^+ = \chi_i + \chi_j$	$9\epsilon(\eta - \delta \sin^3 \theta_e) \sin^3 \theta_e$	<i>Torsión χ simétrica</i>
$\chi^- = \chi_i - \chi_j$	$9\epsilon(\eta + \delta \sin^3 \theta_e) \sin^3 \theta_e$	<i>Torsión χ antisimétrica</i>

Cuadro 1: Constantes de fuerza asociadas a las variables que hacen diagonal la matriz de derivadas segundas del potencial de interacción en el equilibrio, y los tipos de movimientos asociados a estas constantes de fuerza.

Los resultados obtenidos para las constantes de fuerza son correctos, dado que coinciden de manera exacta con las constantes de fuerza que han sido calculadas por el grupo de investigación que me ha guiado en la realización de este trabajo.

8. Conclusiones

Abstract. *In this chapter, we summarize everything we have studied and we have explained, such as the viruses fundamentals, the geometrical properties of icosahedral viruses, the Caspar and Klug's model, electrostatics and mechanical properties of virus, thermodynamics and kinetics of assembly, models of interaction potentials or the paths we took to calculate the force constants starting from our interaction potential. Finally, we comment some applications of the studies made on virology in other research fields or technology.*

La virología física es un campo de la investigación bastante complicado pero importante. Como hemos visto, se puede estudiar muchas cosas sobre estas pequeñas entidades biológicas y sus cápsides. Así, hemos visto la termodinámica y la cinética del autoensamblaje de las cápsides virales, la interacción electrostática que tiene lugar entre las subunidades proteicas que lo conforman, o el estudio de la rigidez que posee la cápside viral, mediante técnicas novedosas que ayudan al desarrollo de la investigación en este campo, como es el caso del microscopio de fuerza atómica (AFM). Por otro lado, hemos estudiado la estructura de los virus icosaédricos, que es la estructura más abundante entre los virus, y, además, han sido descritos y clasificados gracias al trabajo de Caspar y Klug y su número de triangulación T , lo cual ha sido de gran importancia en el campo de la virología.

Utilizando un modelo de potencial de interacción binaria entre capsómeros de tipo *Coarse-grained*, hemos sido capaces de describir la interacción binaria entre trímeros con un desarrollo matemático, a partir del cual hemos podido obtener las constantes de fuerza asociadas a las variables que describen la orientación entre dos trímeros. Para ello, hemos aplicado las condiciones de equilibrio que describen la orientación privilegiada entre dos trímeros necesaria para que se forme la cápside viral.

Gracias a los estudios que se están realizando en virología, podemos comprender estos pequeños cuerpos biológicos, y, así, poder emplearlos a nuestro favor. Este es el caso de su uso en medicina como vehículo de transporte de sustancias en vías de difícil acceso para otros métodos o como una alternativa a antibióticos, o para uso genético, donde se emplea para terapias genéticas [5], o también, en su uso tecnológico, o, simplemente, para poder tratar y eliminar enfermedades desarrolladas por virus, como es el caso del VIH o del Ébola.

Referencias

- [1] M. Požar and R. Podgornik, “Physical virology.” <http://www-f1.ijs.si/~rudi/sola/Physicalvirology%28M.Pozar%29.pdf>.
- [2] J. D. Perlmutter and M. F. Hagan, “Mechanisms of virus assembly,” *Annual Review of Physical Chemistry*, vol. 66, no. 1, pp. 217–239, 2015. PMID: 25532951.
- [3] J. M. G. Llorente, J. Hernández-Rojas, and J. Bretón, “A minimal representation of the self-assembly of virus capsids,” *Soft Matter*, vol. 10, pp. 3560–3569, 2014.
- [4] G. Polles, G. Indelicato, R. Potestio, P. Cermelli, R. Twarock, and C. Micheletti, “Mechanical and assembly units of viral capsids identified via quasi-rigid domain decomposition,” *PLoS computational biology*, vol. 9, p. e1003331, 11 2013.
- [5] A. Luque Santolaria, *Structure, Mechanical Properties, and Self-Assembly of Viral Capsids*. PhD thesis, Universitat de Barcelona, 2011.
- [6] F. H. C. Crick and J. D. Watson, “Structure of small viruses,” *Nature*, vol. 177, no. 4506, pp. 473–475, 1956.
- [7] D. L. D. Caspar and A. Klug, “Physical principles in the construction of regular viruses,” *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 27, pp. 1–24, 1962.
- [8] H. S. M. Coxeter, *Introduction to geometry*. Wiley, 2nd ed., 1989.
- [9] R. Kerner, “Classification and evolutionary trends of icosahedral viral capsids,” *Computational and Mathematical Methods in Medicine*, vol. 9, 09 2008.
- [10] W. H. Roos, R. Bruinsma, and G. J. L. Wuite, “Physical virology,” *Nature Physics*, vol. 6, pp. 733 EP –, Oct 2010. Review Article.
- [11] S. S., *Statistical Thermodynamics of Surfaces, Interfaces, and Membranes*. Addison-Wesley Pub., 1994.

- [12] S. Singh and A. Zlotnick, “Observed hysteresis of virus capsid disassembly is implicit in kinetic models of assembly,” *The Journal of biological chemistry*, vol. 278, pp. 18249–55, 06 2003.
- [13] S. Whitelam, C. Rogers, A. Pasqua, C. Paavola, J. Trent, and P. L. Geissler, “The impact of conformational fluctuations on self-assembly: Cooperative aggregation of archaeal chaperonin proteins,” *Nano Letters*, vol. 9, no. 1, pp. 292–297, 2009. PMID: 19072304.
- [14] S. Whitelam and R. L. Jack, “The statistical mechanics of dynamic pathways to self-assembly,” *Annual Review of Physical Chemistry*, vol. 66, no. 1, pp. 143–163, 2015. PMID: 25493714.
- [15] M. F. Hagan and D. Chandler, “Dynamic pathways for viral capsid assembly,” *Biophysical Journal*, vol. 91, pp. 42–54, jul 2006.
- [16] H. D. Nguyen, V. S. Reddy, and C. L. Brooks, “Deciphering the kinetic mechanism of spontaneous self-assembly of icosahedral capsids,” *Nano Letters*, vol. 7, no. 2, pp. 338–344, 2007. PMID: 17297998.
- [17] M. Hagan and O. Elrad, “Understanding the concentration dependence of viral capsid assembly kinetics—the origin of the lag time and identifying the critical nucleus size,” *Biophysical journal*, vol. 98, pp. 1065–74, 03 2010.
- [18] M. Hagan, “Modeling viral capsid assembly,” *Advances in chemical physics*, vol. 155, 01 2013.
- [19] P. Sorger, P. Stockley, and S. Harrison, “Structure and assembly of turnip crinkle virus: II. mechanism of reassembly in vitro,” *Journal of Molecular Biology*, vol. 191, no. 4, pp. 639 – 658, 1986.
- [20] P. Ceres and A. Zlotnick, “Weak proteinprotein interactions are sufficient to drive assembly of hepatitis b virus capsids,” *Biochemistry*, vol. 41, no. 39, pp. 11525–11531, 2002.
- [21] S. Kasas and G. Dietler, “Probing nanomechanical properties from biomolecules to living cells,” *PFLUGERS ARCHIV-EUROPEAN JOURNAL OF PHYSIOLOGY*, vol. 456, pp. 13–27, 2008. National Licences.
- [22] D. Smith, S. Tans, S. Smith, S. Grimes, D. Anderson, and C. Bustamante, “The bacteriophage ϕ 29 portal motor can package dna against a large internal force,” *Nature*, vol. 413, pp. 748–52, 11 2001.

- [23] A. Evilevitch, L. Lavelle, C. M. Knobler, E. Raspaud, and W. M. Gelbart, “Osmotic pressure inhibition of dna ejection from phage,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, pp. 9292–9295, Aug 2003. 12881484[pmid].
- [24] S. John, *Many-Body Coarse-Grained Interactions using Gaussian Approximation Potentials*. PhD thesis, University of Cambridge, 2016.
- [25] M. Levitt and A. Warshel, “Computer simulation of protein folding,” *Nature*, vol. 253, no. 5494, pp. 694–698, 1975.
- [26] M. A. Padrón, “Autoensamblaje de cápsides virales: Dinámica de cápsides virales icosaédricas: Análisis de modos normales de baja frecuencia,” Master’s thesis, Universidad de La Laguna, 2015.

A. Cálculo de las constantes de fuerza

En primer lugar, vamos a calcular la primeras derivadas del potencial de interacción.

- Con respecto a r :

$$\begin{aligned} \frac{\partial V}{\partial r} = \varepsilon \cdot \left\{ -2m \left(\frac{r_e}{r} \right)^{2m} \frac{1}{r} + 2m \left(\frac{r_e}{r} \right)^m \frac{1}{r} \cdot \left\{ 1 \right. \right. \\ \left. \left. + \alpha \cdot [\sin \theta_i \sin \theta_j \cos(\phi_i - \phi_j) - \cos \theta_i \cos \theta_j - 1] \right. \right. \\ \left. - \frac{\gamma}{2} [\cos^2 \theta_i + \cos^2 \theta_j + 2 \cos^2 \theta_e + 2 \cos \theta_e (\cos \theta_i - \cos \theta_j)] \right. \\ \left. + \zeta [\cos \theta_i \cos \theta_j - \cos \theta_i \cos \theta_e + \cos \theta_j \cos \theta_e - \cos^2 \theta_e] \right. \\ \left. + \eta [-\sin^3 \theta_i \cos 3\chi_i - \sin^3 \theta_i + \sin^3 \theta_j \cos 3\chi_j - \sin^3 \theta_j] \right. \\ \left. - \delta \cdot [\sin^3 \theta_i \sin^3 \theta_j \sin 3\chi_i \sin 3\chi_j] \right\} \end{aligned} \quad (14)$$

- Con respecto a $\phi_i - \phi_j$:

$$\frac{\partial V}{\partial(\phi_i - \phi_j)} = 2\varepsilon \left(\frac{r_e}{r} \right)^m \cdot \left\{ \alpha \cdot [\sin \theta_i \sin \theta_j \sin(\phi_i - \phi_j)] \right\} \quad (15)$$

- Con respecto a θ_i :

$$\begin{aligned} \frac{\partial V}{\partial \theta_i} = -2\varepsilon \left(\frac{r_e}{r} \right)^m \left\{ \alpha [\cos \theta_i \sin \theta_j \cos(\phi_i - \phi_j) + \sin \theta_i \cos \theta_j] \right. \\ \left. - \frac{\gamma}{2} [-2 \cos \theta_i \sin \theta_i - 2 \cos \theta_e \sin \theta_i] + \zeta [-\sin \theta_i \cos \theta_j + \sin \theta_i \cos \theta_e] \right. \\ \left. + \eta [-3 \sin^2 \theta_i \cos \theta_i \cos 3\chi_i - 3 \sin^2 \theta_i \cos \theta_i] - \delta [3 \sin^2 \theta_i \cos \theta_i \sin^3 \theta_j \sin 3\chi_i \sin 3\chi_j] \right\} \end{aligned} \quad (16)$$

- Con respecto a θ_j :

$$\begin{aligned} \frac{\partial V}{\partial \theta_j} = -2\varepsilon \left(\frac{r_e}{r} \right)^m \left\{ \alpha [\cos \theta_j \sin \theta_i \cos(\phi_i - \phi_j) + \sin \theta_j \cos \theta_i] \right. \\ \left. - \frac{\gamma}{2} [-2 \cos \theta_j \sin \theta_j + 2 \cos \theta_e \sin \theta_j] + \zeta [-\sin \theta_j \cos \theta_i - \sin \theta_j \cos \theta_e] \right. \\ \left. + \eta [3 \sin^2 \theta_j \cos \theta_j \cos 3\chi_j - 3 \sin^2 \theta_j \cos \theta_j] - \delta [3 \sin^2 \theta_j \cos \theta_j \sin^3 \theta_i \sin 3\chi_i \sin 3\chi_j] \right\} \end{aligned} \quad (17)$$

- Con respecto a χ_i :

$$\frac{\partial V}{\partial \chi_i} = -2\varepsilon \left(\frac{r_e}{r}\right)^m \left\{ \eta[3 \sin 3\chi_i \sin^3 \theta_i] - \delta[3 \sin 3\chi_j \sin^3 \theta_i \sin^3 \theta_j \cos 3\chi_i] \right\} \quad (18)$$

- Con respecto a χ_j :

$$\frac{\partial V}{\partial \chi_j} = -2\varepsilon \left(\frac{r_e}{r}\right)^m \left\{ -\eta[3 \sin 3\chi_j \sin^3 \theta_j] - \delta[3 \sin 3\chi_i \sin^3 \theta_i \sin^3 \theta_j \cos 3\chi_j] \right\} \quad (19)$$

Teniendo las primeras derivadas, vamos a derivar cada una de ellas por cada variable que caracteriza nuestro problema. De esta manera, vamos a obtener una matriz de segundas derivadas. Dado la gran extensión que tendría esta matriz de segundas derivadas, y la gran cantidad de términos que resultaría de ella, vamos a dejar para el lector, si lo desea, este cálculo numérico. Una vez obtenida esta matriz de derivadas segundas, vamos a imponerle las condiciones de equilibrio de nuestro problema, descritos en el apartado 7.2. Tras realizar esta operación, nos queda la siguiente matriz en función de r , $\phi_i - \phi_j$, θ_i , θ_j , χ_i y χ_j , en ese orden:

$$\begin{pmatrix} 2\varepsilon m^2 \frac{1}{r_e^2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 2\varepsilon \alpha \sin^2 \theta_e & 0 & 0 & 0 & 0 \\ 0 & 0 & 2\varepsilon(\alpha + \gamma \sin^2 \theta_e) & 2\varepsilon(\alpha - \zeta \sin^2 \theta_e) & 0 & 0 \\ 0 & 0 & 2\varepsilon(\alpha - \zeta \sin^2 \theta_e) & 2\varepsilon(\alpha + \gamma \sin^2 \theta_e) & 0 & 0 \\ 0 & 0 & 0 & 0 & -18\varepsilon \eta \sin^3 \theta_e & -18\varepsilon \delta \sin^6 \theta_e \\ 0 & 0 & 0 & 0 & -18\varepsilon \delta \sin^6 \theta_e & -18\varepsilon \eta \sin^3 \theta_e \end{pmatrix}$$

Como podemos observar, esta matriz no es diagonal. Sin embargo, es diagonal por bloques, por lo que tenemos el bloque ' θ ' y el bloque ' χ '. Si diagonalizamos la matriz de cada bloque, nos queda lo siguiente:

- Bloque ' θ ':

$$\begin{vmatrix} 2\varepsilon(\alpha + \gamma \sin^2 \theta_e) - \lambda & 2\varepsilon(\alpha - \zeta \sin^2 \theta_e) \\ 2\varepsilon(\alpha - \zeta \sin^2 \theta_e) & 2\varepsilon(\alpha + \gamma \sin^2 \theta_e) - \lambda \end{vmatrix} = 0$$

Diagonalizando este determinante, nos quedan lo siguientes autovalores:

$$\lambda_{\theta+} = 2\varepsilon(2\alpha + (\gamma - \zeta) \sin^2 \theta_e) \quad \lambda_{\theta-} = 2\varepsilon(\gamma + \zeta) \sin^2 \theta_e$$

Como podemos observar, nos ha resultado el doble de las constantes de fuerza de θ^+ y θ^- , respectivamente. Esto es debido a que, si obtenemos los vectores asociados a esos autovalores, nos salen los siguientes vectores:

$$\theta^{+'} = \frac{1}{\sqrt{2}}(\theta_i + \theta_j) \quad \theta^{-'} = \frac{1}{\sqrt{2}}(\theta_i - \theta_j)$$

Es decir, los vectores asociados a estos autovalores van multiplicados por el factor $\frac{1}{\sqrt{2}}$. Para obtener las variables deseadas, como se trata de una forma cuadrática, si multiplicamos los estados $\theta^{+'}$ y $\theta^{-'}$ por un factor $\sqrt{2}$, los autovalores se dividirán por un factor 2, quedando como resultado, las variables y las constantes de fuerza que aparecen en la tabla 1.

- Bloque ' χ ':

$$\begin{vmatrix} -18\varepsilon\eta \sin^3 \theta_e - \lambda & -18\varepsilon\delta \sin^6 \theta_e \\ -18\varepsilon\delta \sin^6 \theta_e & -18\varepsilon\eta \sin^3 \theta_e - \lambda \end{vmatrix} = 0$$

Diagonalizando este determinante, nos quedan lo siguientes autovalores:

$$\lambda_{\chi^+} = 18\varepsilon(\eta - \delta \sin^3 \theta_e) \sin^3 \theta_e \quad \lambda_{\chi^-} = 18\varepsilon(\eta + \delta \sin^3 \theta_e) \sin^3 \theta_e$$

Como podemos observar, nos queda el mismo caso que en el bloque ' θ '. Por lo que, análogamente al bloque ' θ ', nos queda como resultado las variables y las constantes de fuerza que aparecen en la tabla 1.

Si escribimos la matriz diagonal con las nuevas variables, nos queda la matriz de constantes de fuerza:

$$\begin{array}{cccccc} 2\varepsilon m^2 \frac{1}{r_e^2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 2\varepsilon\alpha \sin^2 \theta_e & 0 & 0 & 0 & 0 \\ 0 & 0 & \varepsilon(2\alpha + (\gamma - \zeta) \sin^2 \theta_e) & 0 & 0 & 0 \\ 0 & 0 & 0 & \varepsilon(\gamma + \zeta) \sin^2 \theta_e & 0 & 0 \\ 0 & 0 & 0 & 0 & 9\varepsilon(\eta - \delta \sin^3 \theta_e) \sin^3 \theta_e & 0 \\ 0 & 0 & 0 & 0 & 0 & 9\varepsilon(\eta + \delta \sin^3 \theta_e) \sin^3 \theta_e \end{array}$$