
# Statistical mechanical models of virus capsid assembly

by Stephen Daniel Hicks

---

# STATISTICAL MECHANICAL MODELS OF VIRUS CAPSID ASSEMBLY

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Stephen Daniel Hicks

August 2010

STATISTICAL MECHANICAL MODELS OF VIRUS CAPSID ASSEMBLY

Stephen Daniel Hicks, Ph.D.

Cornell University 2010

Viruses have become an increasingly popular subject of physics investigation, particularly in the last decade. Advances in imaging of virus capsids—the protective protein shells—in a wide variety of stages of assembly have encouraged physical assembly models at a similarly wide variety of scales, while the apparent simplicity of the capsid system—typically, many identical units assembling spontaneously into an icosahedrally symmetric (rather than amorphous) shell—makes the problem particularly interesting. We take a look at the existing physical assembly models in light of the question of how a particular assembly target can be consistently achieved in the presence of so many possible incorrect results. This review leads us to pose our own model of fully irreversible virus assembly, which we study in depth using a large ensemble of simulated assembled capsids, generated under a variety of capsid shell elastic parameters. While this irreversible model (predictably) did not yield consistently symmetric results, we do glean some insight into the effect of elasticity on growth, as well as an understanding of common failure modes. In particular, we found that (i) capsid size depends strongly on the spontaneous curvature and weakly on the ratio of bending to stretching elastic stiffnesses, (ii) the probability of successful capsid completion decays exponentially with capsid size, and (iii) the degree of localization of Gaussian curvature depends heavily on the ratio of elastic stiffnesses. We then go on to consider more thoroughly the nature of the ensem-

ble of symmetric and almost-symmetric capsids—ultimately computing a phase diagram of minimum-energy capsids as a function of the two above-mentioned elastic parameters—and also look at a number of modifications we can make to our irreversible model, finally putting forth a rather different type of model potentially appropriate for understanding immature HIV assembly, and concluding with a fit of this new model's parameters to recent experimental structures.

A common thread between the coarse-grained models we discuss in the first part of the thesis is that they all depend explicitly on elastic parameters that are otherwise completely unmotivated. We thus devote the second part to the question of how (elastic) model parameters can be determined from *ab initio* methods. Modeling protein interactions as springs with very general quadratic potentials, we run atomistic molecular dynamics simulations and analyze the trajectories to determine stiffness tensors for these generalized springs. After a thorough examination of the mathematical structure of our springs—including transformations of the stiffness tensors into different reference frames and gauges, and an analytical formula for composing generalized springs in series—we go on to apply the technique to measure the elasticity of a mature HIV capsid lattice by simulating isolated pairs of interacting protein domains. We compute the relaxation times for each bond, and for the entire lattice, which both gives the stiffness as a physical, comparable timescale, and also provides a way to invalidate many simulations with too short a run time. Because calculation of the relaxation matrix requires a measurement of the diffusion of the individual proteins, we conclude with a brief study of the effects of finite box sizes and differing thermostat strengths on diffusion measurements from atomistic simulations.

## BIOGRAPHICAL SKETCH

In 1981, Stephen Daniel Hicks was born in Tampa, Florida, in whose suburbs he proceeded to spend eighteen formative years of his life. His love of the natural sciences developed at an early age, with the support of parents who provided him everything from grapes (for counting, of course) and algebra to slide rules and computers. After exhausting his resources at home, he continued his studies in high school by spending much of his free time training for and competing in mathematics contests, and by participating in a sequence of pre-engineering courses, which landed him a software engineering internship at the Tampa Electric Company in the spring and summer of 1999.

Stephen moved to Gainesville to attend the University of Florida in the fall of 1999, intending to major in Mathematics and Computer Science. During his first summer he took on a research project under Professor Meera Sitharam studying equiseparations—a construction ultimately related to the P vs. NP problem in a particular restricted computational environment. After finishing the honors physics sequence, Stephen decided to change majors and began studying Physics instead of Computer Science. In the summer of 2001, he participated in a summer research program at the Ohio State University where he did experimental work on organic electronics, supervised by Professor Art Epstein. During his senior year, he studied the effect of various (sometimes esoteric) Fermi surfaces on resulting magnetoresistance under Professor Selman Hershfield. In addition to research, he participated in physics outreach events with the Society of Physics Students, and was involved in residence hall government. He graduated summa cum laude in May 2003 and in the fall enrolled in a Physics Ph.D. program at Cornell University in Ithaca, New York.

In the summer of 2004 Stephen began his thesis research on virus assembly, supervised by Professor Christopher Henley. During his graduate career, he spent several semesters as a teaching assistant for a variety of physics courses, and continued to do physics outreach with the Cornell Center for Materials Research outreach program. While in Ithaca, Stephen's free time was divided between Bible studies, board games, martial arts, whitewater kayaking, and computer programming, the latter of which developed into a business in 2007 when Professor Andy Ruina hired him part-time to develop and maintain some LaTeX macros. After winning the Judges' Prize in the International Conference for Functional Programming annual contest in 2008, he began to consider pursuing a career in software engineering. In January 2010, he took a programming job at Google, Inc., and currently lives in Mountain View, California, where he works on software for monitoring and optimizing datacenter power utilization efficiency.

*For Mom and Dad, who taught me to stop and smell the roses.*

**ACKNOWLEDGEMENTS**

A Ph.D. is as much the work of a single individual as the raising of a child is the work of the parents alone. Indeed, I'm indebted to many people, without whose guidance, advice, admonition, support, and encouragement this thesis would never have come together. Moreover, In attempting an enumeration of these acknowledgments, I've certainly left out a few people who deserve inclusion.

First I'm grateful to my adviser, Professor Chris Henley, for opening my eyes to the unending myriad of universes available to study in condensed matter physics, and for guiding me along my own exploration of one such universe. Under his guidance I've learned not only a variety of techniques for approaching difficult scientific problems, but also how to communicate the results precisely and effectively, and with a subtle haiku or pun thrown in for good measure.

A number of others on the faculty both in the physics department and outside have been extremely helpful as well: my committee, Professors Veit Elser, Rob Thorne, and Volker Vogt, helped me to refine my research and the presentation of it, and also provided me some interesting problems to think about on the side. It's also important to have support from outside the committee, and I'm grateful to Professor Jim Sethna for advice on several occasions, and to Professor Andy Ruina for exercising my hobby which eventually grew into a career. Further, I would be remiss if I neglected my undergraduate mentor, Professor Selman Hershfield, and some particularly encouraging teachers: Professor Jonathan King, Susan Hammer, Terry Adams, and Ervin Nevismal.

I'm also thankful to have been part of a large and social class of fellow graduate students, who made late nights working on E&M problem sets not quite so painful; and particularly to Kelken Chang and Leah Shaw, who on several occasions sent me

useful references even after they had left Ithaca long behind.

While sound academic support was crucial to completing a project of this magnitude, I am equally appreciative of the social and spiritual support I received. The congregation at Christ Chapel—particularly Pastor Steven Felker, Lawrence Lyon, and the choir—went a long way in looking after my spiritual well-being. Outside of church, the Graduate Christian Fellowship provided me with spiritual growth through challenging Bible studies, and more importantly, a rich set of friends to count on for just about anything I could need: in particular Ray Fertig III, Geoff Recktenwald, and David Roundy, with each of whom I've talked for hours at a time about everything from physics to theology to relationships. Without such friends, I could never have kept my sanity for six years.

From a much greater distance, I'm thoroughly grateful to my parents, who throughout my brother's and my childhood always found a healthy balance of pushing us to persevere and succeed, while loving us unconditionally regardless of our failures. I'm also grateful for the love of nature and the thirst for knowledge that they instilled in us at all stages of our lives, and for continuing to support and advise me even a full decade after leaving the nest.

Finally, none of this would have been remotely possible but for the abundant grace of God, who has straightened my path every time I've begun to stray. I'm thankful that He has blessed me with amazing friends and family, with the privilege of spending twenty three years of my life in school, and with the capacity to continue to learn throughout. Even more, I'm thankful for my redemption, achieved on the cross, and I pray that this work is a worthy response to that outpoured grace.

**TABLE OF CONTENTS**

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| aa | Amino acid(s) |
| AFM | Atomic force microscopy |
| ATP | Adenosine triphosphate |
| BPTI | Bovine pancreatic trypsin inhibitor |
| CA | Capsid domain (of retroviral Gag) |
| CCMV | Cowpea chlorotic mottle virus |
| CHARMM | Chemistry at HARvard Macromolecular Mechanics |
| CK | Caspar–Klug |
| cryo-EM | Electron cryomicroscopy |
| CTD | C-terminal domain (of retroviral CA) |
| DNA | Deoxyribonucleic acid |
| ED | Essential dynamics |
| EM | Electron microscopy |
| ESCRT | Endosomal sorting complex required for transport |
| HIV | Human immunodeficiency virus |
| $IP_6$ | Inositol phosphate |
| MA | Matrix domain (of retroviral Gag) |
| MBD | Membrane binding domain (of retroviral MA) |
| MD | Molecular dynamics |
| MLV | Murine leukemia virus |
| NAMD | NAnoscale molecular dynamics (software) |
| NC | Nucleocapsid domain (of retroviral Gag) |
| NMA | Normal mode analysis |
| NMR | Nuclear magnetic resonance |
| NTD | N-terminal domain (of retroviral CA) |
| PCA | Principal component analysis |
| RNA | Ribonucleic acid |
| RSV | Rous sarcoma virus |
| SP | Spacer region (of retroviral Gag) |

## LIST OF SYMBOLS

| | |
|---|---|
| $\mathsf{\pi}$, $\mathsf{i}$, $\mathsf{e}$ | Constants are upright |
| $r$, $\phi$, $\theta$ | Scalar variables are italic |
| $\boldsymbol{x}$, $\boldsymbol{\theta}$ | Vectors are bold italic |
| $\mathbf{\Omega}$, $\mathbf{\omega}$, $\mathbf{K}$ | Matrices and tensors are bold upright |
| $\mathbb{R}$, $\mathbb{r}$ | Rigid body coordinates are blackboard bold |
| $\mathfrak{q} = \rho + \mathsf{i}\boldsymbol{\zeta}$ | Quaternions are fraktur |

CHAPTER 1

# INTRODUCTION

In recent years the question of spontaneous assembly has arisen in many apparently disconnected fields, including nanofabrication [1, 2], robotics and microelectronics [3, 4], and particularly biology [5, 6]. While the assembly of many biological structures, such as actin filaments and chromatin, requires energy in the form of ATP hydrolysis [7, 8, 9], numerous other structures assemble spontaneously. In particular are lipid bilayers [10, 11] and many virus capsids [12, 13]. In this dissertation, we look at several models of virus capsids. While our ultimate goal is to describe the assembly process, we digress as well to mechanical models as a means of understanding the parameters input into the assembly models.

## 1.1 Outline

In the following chapters we will explore some physical aspects of virus assembly in more detail, starting with top-down models in Chapters 2–4, and then concluding with bottom-up experiments in Chapters 5–7.

In Chapter 1, we provide an introduction to the biological system of retroviruses that we generally have in mind during the rest of the thesis. Following this is a literature review of pertinent experiments in retrovirus assembly, of physical models of virus assembly, and finally of important atomistic computational techniques.

Chapter 2 is taken primarily from a *Physical Review E* paper [14] in which we describe an irreversible growth model for capsid assembly. In Chapter 3, we look deeper at the energetics of this model and the ensemble of capsids one can build. We then move on to the dynamics in Chapter 4, exploring different parametrizations of capsids, and along with this, different growth rules for transitioning between partially

assembled capsids. We conclude this chapter by proposing a model of immature human immunodeficiency virus (HIV) assembly.

Seeking to provide a microscopic motivation for these continuum models, in Chapter 5 we explore a formalism of generalized springs, whose stiffnesses can be calibrated by molecular dynamics (MD) data. Chapter 6 demonstrates the use of this technique with the mature HIV capsid and is taken from a *Physical Review E* rapid communication [15]. Finally, in Chapter 7 we explore an interesting question that arose from the MD analysis: what effect do the thermostat and the finite size of the simulation cell have on measurements of diffusion, and can we make use of these effects to efficiently measure diffusion with MD simulations?

## 1.2 Virus biology and structure

Viruses occupy a peculiar place in biology. Most simply, a virus consists of two major components: the genome and the capsid. The genome is a piece of nucleic acid (single-stranded DNA, double-stranded DNA, or RNA) that codes for the viral proteins. While many viruses contain a number of auxiliary proteins (such as proteases or reverse transcriptases), each virus has at least one type of capsid protein (and most have only one—notable exceptions are the "pseudo" $T$-number viruses, such as the pseudo $T = 3$ rhinoviruses). Depending on the species, anywhere from sixty to thousands of copies of this capsid protein assemble with a copy of the genome inside (whether nucleated around the genome as in retroviruses, or stochastically including parts of the genome as in cowpea chlorotic mottle virus (CCMV), or pumping in the genome after assembly finishes, as in bacteriophages). An effective working definition is that a virus is a protein coat surrounding a viral genome.

Viruses come in a wide variety of shapes and sizes, ranging from sixty to thou-

sands of proteins. In addition, most known capsids have icosahedral or cylindrical point group symmetry [16][1], forming spheres and icosahedra in the former case, and tubes in the latter[2]; in both cases, the ensemble of assembled capsids for a given species is monodisperse (under fixed environmental conditions). However, a number of species form irregular capsids, which, by their asymmetry, are almost necessarily polydisperse, or pleiomorphic. Most irregular capsids are roughly spherical, though some are more exotic: murine leukemia virus (MLV) forms tubes, and (of particular interest recently) are the unique conical capsids of HIV. Finally, a very few viruses form regular capsids that are less than maximally symmetric, i.e. $\phi$29's prolate capsid, which has $T = 3$ end caps connected by a $T = 7$ midsection. It is unknown whether assembly of regular viruses is correct most of the time, or if a large number of incorrectly-assembled capsids are somehow digested (particularly if the capsid has holes) or discarded before the micrographs are averaged.

## 1.3 Retroviruses

Retroviruses, such as HIV and rous sarcoma virus (RSV), are RNA viruses that all contain a characteristic enzyme—reverse transcriptase—allowing the RNA to be reverse-transcribed into DNA and inserted into the infected cell's nucleus. The infected cell produces many copies of the viral proteins (in particular, the structural polyprotein, Gag) and genome, which then aggregate at and assemble on the cell membrane be-

[1] It was noted early that capsids ought to be built from many identical copies of comparatively small proteins, in order to maximize the volume available for the genome, while minimizing the space on the genome needed to code them [17]. However, point group symmetries set an upper limit of 60 units which can be joined equivalently to form a closed convex polyhedron [17, 18]; thus many proteins must be inequivalent by the symmetry

[2] Our models will also admit flat sheets, but these are obviously not biologically viable since there is no way to package the genome.

fore budding out of the cell as an immature virus particle. A maturation step then takes place in which a protease cleaves Gag into its primary constituent proteins: matrix (MA), capsid (CA), and nucleocapsid (NC). The CA domain is further composed of a dimer-forming C-terminal domain (CTD) and a hexamer-forming N-terminal domain (NTD), connected by a short peptide linker[3]. After cleavage, the MA remain bound to the lipid membrane, the NC remain bound to the RNA, and some or all of the CA end up in the mature capsid. With so many steps in the virus lifecycle, there are a number of things to investigate: we will focus primarily on the actual assembly of the immature and mature particles, with a tabulated summary of key results in Table 1.1.

### 1.3.1 Assembly *in vitro*

Campbell and Vogt [26] pioneered the process of *in vitro* assembly of RSV from fragments of Gag in the presence of RNA. The requirement of the nucleic acid led to a proposed model of RNA-nucleated assembly, which was further developed by replacing the RNA-binding zinc fingers of the NC domain with a leucine zipper [31][4]. Ganser et al. [27] extended this work to HIV, using electron cryomicroscopy (cryo-EM) to visualize the cones characteristic of mature HIV, modeled as fullerene cones with the correct discrete cone angles [35]. Li et al. [36] purified HIV tubes, allowing a measure-

[3]Each retrovirus has a variety of different spacer regions in addition to the three primary domains, such as p1, p2, and p6 in HIV (reviewed in [19, 20, 21]), and SP and p10 in RSV.

[4]The leucine zipper effectively adds an artificial dimerization interface between pairs of proteins. The fact that this rescued assembly is thought to indicate the importance of these Gag-Gag interactions (and in particular, dimerization) mediated by the RNA through the zinc fingers (also known as Cis-His boxes) of the NC domain. Because the CTD interaction is also dimeric, it is thought that the RNA-mediated interaction may cause a conformational change to somehow make the proteins competent to polymerize.

Table 1.1: Summary of some known retroviral assembly morphologies

| domains | virus | requirements | morphology |
|---|---|---|---|
| CA only | HIV | high salt | tubes |
| " | RSV | acid pH | tubes, spheres[a] |
| " | " | high $PO_3$ | $T = 1$, $T = 3$ [22] |
| " | " | CA-SP[b] | " [23, 24] |
| " | " | protonate $D_{191}$ | " [25] |
| CA-NC | HIV,RSV | RNA | tubes, cones[c] [26] |
| " | HIV | high salt | tubes, cones [27] |
| ΔMA Gag[d] | " | RNA | spheres[a] |
| Full Gag[e] | " | RNA, $IP_6$ | " [28][f] |
| ΔMBD Gag | RSV | RNA | " [29] |
| Delete NC | HIV,RSV | *in vivo* | no assembly |
| Delete MA | HIV | " | spheres[a] [30] |
| NC → zipper | HIV | " | " [31] |
| Break SP helix[g] | RSV | " | tubes [32] |
| Unmyristoylated[h] | HIV | " | no budding |
| Break late domain | HIV,RSV | " | " |
| Full Gag | " | " | spheres[a,i] |

[a] except where $T$-numbers noted, spheres are irregular
[b] only a small (non-stoichiometric) amount needed
[c] the cones are only found from HIV
[d] the MA domain in this construct is missing all but the last 20aa
[e] this construct unmyristoylated and missing (C-terminal) p6 domain
[f] in the absence of $IP_6$, the spheres were significantly too small
[g] the putative SP helix is discussed in [33]
[h] removing the N-terminal Gly residue prevents myristoylation
[i] it is believed that *in vivo* assembly can occur even with no RNA [34]

ment of the properties of the CA lattice. Later studies used tomographic techniques to go beyond the two-dimensional cryo-EM images and produce three-dimensional reconstructions of single virus particles [37, 38]. There is hope that such better pictures will help to answer the question of which end of the cone forms first.

These pictures of mature assembled capsids revealed a wide polymorphism in the results: HIV naturally forms a variety of tubes and cones [39] with diameters around 140nm, while RSV produces irregular sphere-like capsids—sometimes described as coffins and lozenges [40, 41]—with a distribution in the degree of asphericity [42] and diameters around 80nm. In addition to the natural pleiomorphism, mature HIV can assemble spheres, tubes, or cones depending on protein mutations [43] and even on environmental conditions such as pH [44, 45]. Immature particles of both RSV and HIV form spheres of different diameters [46], and it was found by combining different parts of RSV and HIV together into chimeras that CA is the determining factor of the particle size [47].

Several important morphologies were later discovered by experimenting with mutations and environmental conditions. One particular mutation in HIV was found to produce very large capsids, whose surfaces were nearly flat [48], leading to high-resolution cryo-EM averaging and eventually to crystallization of an HIV hexamer [49]. On the other end of the spectrum, the presence of high concentrations of divalent anions was found to produce small icosahedral $T = 1$ and $T = 3$ capsids in RSV [22]. Similar morphologies can also result from dimerization catalyzed by a small amount of CA-SP [23, 24] or by acidification of a certain residue [25]. Inositol phosphate ($IP_6$) has been shown to change the structure of Gag significantly, resulting in very small virus-like particles [28, 50].

### 1.3.2 Assembly *in vivo*

While *in vitro* assembly can produce capsids resembling either mature or immature particles, assembly in live cells must proceed along the biological pathway, and thus *in vivo* assembly produces only immature particles (although if the particles can exit the cell and have a working protease, they will mature).

In general, viruses exit the host cell by either inducing cell death or by budding out, the latter of which can occur either during or after assembly. Retroviruses are known to assemble the immature capsid on the membrane, such that budding is simultaneous with assembly [51, 52]. It stands to reason therefore that the membrane, and its interaction with the MA domain [53, 54, 55] is important to the final budded morphology.

A large number of experiments have investigated mutations that interfere with retroviral budding[5]. In particular, each retrovirus has one or more late domains—short sequences of about four amino acids—that are critical for budding [58, 59, 60, 61, 62, 63]. The current explanation is that the late domains recruit the cellular ESCRT complexes to pinch off the budded virus, although the exact mechanism is not fully understood [64]. When the late domains are knocked out, the result is a large number of nearly-complete capsids sticking out of the cell, but not pinched off. Similar pinch-off failure has also been observed as a result of mutations to the NC region, far away from the late domains [65].

Keller et al. [32] mutated another region of RSV Gag to find changed budding morphology: while wild type immature RSV is known to be roughly spherical, breaking the (putative) six helix bundle between the CA and NC domains forms tubes, as well

[5]Typically budding is observed with snapshots taken by cryo-EM, but recent advances have opened the door to observation of individual budding events in real-time [56, 57].

as some more exotic “chained bead” structures.

### 1.3.3 Gag-Gag interactions

These studies of the effect of mutations and environment on *in vitro* and *in vivo* assembly, together with improved microscopy, shed some light on the details of the interactions in mature and immature capsids. The mature CA has a dimer interaction in the CTD [66, 67] and a hexamer interaction in the NTD [68, 49]. On the other hand, the interactions in immature Gag are not so well understood. It has been shown that MA trimerizes [69] and can also hexamerize . The NTD hexamer interaction does not appear to be present in wild-type immature capsids [33], but the CTD dimer is likely to be the same across both mature and immature. A domain-swapped dimer has also been observed [70], but is not generally believed to be biologically relevant. The most important non-dimer bond[6] in immature Gag appears to be a six-helix bundle in the spacer region between CA and NC. While NC particles do not interact directly, they are known to be important in nucleating assembly by their RNA-mediated interactions [31]. Finally, it is possible that the different interactions are not strictly separated between mature and immature particles, particularly when mutations are involved: while some pairs of interactions are thought to be allosterically prevented (e.g. the six-helix bundle forces the NTD hexamer binding sites apart [33]), there may be others that can mix and match.

While the interactions cited above are largely specific to HIV (with the exception of the first NTD hexamer crystal structure, which was for MLV [68]), the major Gag domains (MA, CA, and NC) are mostly homologous across different retroviruses: while the residue sequences are different, the secondary and tertiary structures are very

[6] see §4.3 for more on why we look for a “non-dimer” bond.

similar.

In light of all the different interactions, it is useful to distinguish one interaction as "strongest" or "primary" (see in particular §4.3). Although bond energies are available for some viruses [71], we are actually more interested in determining which oligomeric subunits are present in solution as the building blocks for the capsid. In the case of HIV and RSV, the experiments on dimerization [31], as well as gel electrophoresis, have suggested that dimers are first to form in solution, and would thus make the most sense as a unit in an assembly model.

### 1.3.4 Stoichiometry

One simple question that has provoked extensive debate is how many copies of the Gag protein are in immature and mature viruses. In particular, since not all the Gag from the immature capsid ends up in the mature capsid, this must be measured for both versions. It is estimated that around 1500 Gag are in the immature RSV capsid [72, 73]. Measurements of the stoichiometry of immature HIV range from 750–2500 [74], using fluorescence fluctuation spectroscopy, to 5000 [75], by dividing the area of the capsid by the area of a unit cell. Typical mature HIV cores are thought to contain 1200–1500 copies of CA [75, 76].

A major difficulty in this accounting is now understood to be due to the patchy nature of the immature capsid. It was discovered early that immature retrovirus particles had some at least some lattice order [77], but more recent studies have shown that there are in fact holes in the immature capsid [33, 78], leading to incorrect conclusions based on the size of the immature particles.

## 1.4 Physical assembly models

The last decade has seen a surge in physical virus models, the majority of which deal with the process (or the results) of assembly[7]. Broadly speaking, we will look at three categories of models. First are models that ignore all of the dynamics and consider only the energy of fully-assembled capsids—we will call these *equilibrium models*. On the opposite end of the spectrum are dynamical models in which the resulting fully-assembled capsid is predetermined at the outset (such models are typically designed to explore the pathway or the kinetics of assembly)—we will call these *guided assembly models*. Somewhere in between are dynamical models that leave the final result free, which we will call *unguided assembly models*.

Nearly all of these models are based to some degree on the framework of quasiequivalence. After looking briefly at this framework, we will review the important models from each of the above categories in turn.

### 1.4.1 Quasiequivalence

A typical virus uses only one, or a few, kinds of protein in its capsid, with copy numbers in the hundreds or even thousands; consequently, most icosahedral capsids are necessarily built from copies of the sameu nit in positions that are *not* equivalent by any global symmetry, since the order of the icosahedral group is only sixty. However, Caspar and Klug [18] identified an elegant approximate symmetry they called *quasiequivalence*, which has grown into a framework used in most capsid models[8].

[7] Other models have dealt with the post-assembly packing of DNA into phages [79, 80] or budding of complete capsids [81].

[8] The specific meaning of "quasiequivalence" has unfortunately diverged over time and now means anything from an assertion that symmetry-inequivalent units are small variations of the same type of bond (as we understand it), to the exact opposite (in the "local rules" literature), making communication between researchers tricky.

The key idea is that locally, every bit of the capsid is a patch of triangular lattice; in an infinite triangular lattice, all the units *would* be symmetry equivalent. Caspar and Klug argued that typical proteins could accommodate a variation of $\pm 5^\circ$ in bond angles [82], while qualitatively maintaining the same microscopic bonding between proteins. This allows representation of any capsid as a network of approximately equilateral triangles, with a constraint (due to the bond angle limitation) that the number of triangles around every vertex must always be either five or six.

The points of local five-fold symmetry may be identified with the topological defects called *disclinations* (to be defined in §2.1.1), and any closed shell must contain exactly twelve of them[9]. In an icosahedral capsid, the disclinations form the vertices of a large icosahedron, the edges of which have length $\sqrt{T}$ in lattice units, where the triangulation number $T = 1,3,4,7,\ldots$ is one of a sequence of discrete allowed integers [18], so that there are $60T$ small triangles. Some researchers believe that capsids larger than about $T = 7$ do not actually assemble regularly without auxiliary proteins [84].

We emphasize that the rules of quasiequivalence do not force any global symmetry, nor do they fix the size of the completed capsid. Thus it is surprising that many viruses reliably assemble large symmetric capsids. The challenge to theory is to explain both the size and shape selection, or at least to explain why a closed shell is formed, when tubes or sheets would be equally consistent with the local bonding. It would not be surprising if models predict different capsids depending on parameters (which might experimentally correspond to pH, salt content, catalysts, protein concentrations, or mutations in the capsid protein). Such polymorphic behavior is very fruitful to study in quasiequivalent models: it effectively explores more of the

[9]This is true provided the shell has genus 0. A near-exception is the torovirus, whose cylindrical capsid is bent into an open torus [83].

various local geometries in which the proteins can bind and thus can allow more parameters to be determined, in principle. Chapter 2 develops a model of irreversible (non-equilibrium) assembly of quasiequivalent units which produces a highly polymorphic ensemble of capsids, which we argue below may model the growth of retrovirus capsids.

It is worth mentioning that Twarock [85] has developed an alternative framework to Caspar–Klug (CK) quasiequivalence using rhombs and kites rather than triangles, and can therefore describe the anomalous viruses from *Papovaviridae*.[10]

### 1.4.2 Equilibrium models

By completely doing away with the dynamics, the simplest capsid models can focus entirely on answering the question of which capsids are most energetically favorable. In these models, a microscopically motivated phenomenological Hamiltonian is shown to be optimized by certain shapes, and it is assumed that this free energy minimum is found during the actual assembly process. Thus Bruinsma et al. [13, 86] modeled pentamers and hexamers as different-sized discs packed on a sphere, with an effective Hamiltonian favoring dense packing, a bending stiffness with spontaneous curvature, and a switching cost to make pentamers (rather than hexamers) of the proteins. When this switching cost is small, icosahedral viruses were selected over nonicosahedral shapes [86]. Additionally they demonstrated polymorphism, similar to phenomena seen in CCMV, by showing a phase transition between tubes, $T = 3$,

[10] One interesting extension to this theory, devised by Chris Henley, is to treat the kites instead as a combination of an equilateral triangle and a thin triangular "missing wedge". In this case, there are now three units with symmetries that correspond to the units they actually represent: rhombs representing dimers, triangles representing trimers, and wedges to connect the two types of units (since they have different edge lengths). This would provide an interesting starting point for an assembly model targeted at *Papovaviridae*, but we never explored it further.

and $T = 1$ capsids as the model parameters varied. Nguyen et al. [87, 88] have also attempted to understand conical retrovirus capsids with an equilibrium theory, although the assumptions (particularly of fixed numbers of subunits) were not biologically reasonable. Another family of models, introduced by Nelson [89], focuses on the external shape of large capsids, using continuum elastic theory: the shape evolves from practically spherical to sharply faceted as the size increases or the bending stiffness decreases[11].

These models provide an important foundation, but knowing that a particular capsid is energetically favorable does not explain how it was assembled.

### 1.4.3 Guided assembly models

The next category of models goes slightly beyond the equilibrium models by picking out (typically) the energy-minimizing results and constructing a model such that the desired result is the only possible assembly. One family of models in this category are the kinetic models, pioneered by Zlotnick and coworkers [90, 91]. They generally construct $T = 1$ capsids out of twelve pentameric building blocks (alternately, arbitrary $T$-numbers with $5T$-meric building blocks) using chemical rate equations and the Law of Mass Action. In the simplest models [90, 92], only the partial capsids that minimize perimeter (i.e. line tension) are allowed, so that there is a single predetermined pathway, but later work showed that allowing other partial capsids (with longer perimeters) is merely a small perturbation [93]. The chief result of this body of work is the idea of the *kinetic trap*: if the product of the forward assembly rate

[11] These models present a very well-developed elasticity theory, but the researchers do not believe that spontaneous curvature is a biological phenomenon, so they unfortunately lack any work in the direction of understanding the interaction of disclinations with any background curvature.

and the initial concentration of monomers is large, then the monomers are quickly depleted in assembling partial capsids, at which point forward progress slows significantly, effectively waiting for some partial capsids to disassemble in order to free up monomers. Morozov et al. [94] consider the continuum limit of these kinetic theories, resulting in a Fokker-Planck equation to describe assembly. Finally, Keef et al. [95] extend Zlotnick's work with Twarock's tiling theory [85] and consider the effect of different association energies on the kinetics of *Papovaviridae* assembly.

Another family of models in this category is based on the work of Berger et al. [96], under the heading of "local rules". These models are able to assemble larger ($T > 1$) capsids by assuming multiple "conformations" of the capsid protein, with matching rules imposed such that there is a single unique structure that obeys all the matching rules. In the strictest implementation, $T$ different conformations, each with specific geometry and matching rules, are required to build a capsid with a given $T$-number, and errors are only admitted by placing two units at a pair of topologically equivalent locations [96, 97, 98]. It is generally necessary to assume that the same capsid protein molecule has different conformation species, each of which has entirely different specific binding. It appears implausible that so many different functions could be built into one molecule, or that evolution could have discovered this solution, if it is the only way to engineer a large capsid.

A weaker form of local rules, with ambiguous binding specificities, allows fewer than $T$ conformations [99]. Similar results can be achieved without matching rules by specifying $T$ different unit geometries perfectly aligned for the target structure, allowing only one binding site on each unit to be active at a time, and then breaking bonds that don't settle into the correct geometry after a certain amount of time [100].

Because these models essentially have no freedom in their final morphology, they

cannot address the question of how a particular capsid morphology is chosen from among the vast sea possible results, without first giving a compelling explanation for the mechanism behind the matching rules (one possible mechanism is allostery: if the conformational change results in a different bond angle at one location on the unit, as well as a changed binding site elsewhere, then some simple local rules, such as those for $T = 3$ capsids [99], may be feasible). Instead, local rules models are useful for understanding and simulating the kinetics and assembly pathways, while bypassing the question of how the capsid ends up in the correct structure. Recent local rules simulations have suggested the importance of environmental conditions (i.e. concentrations) in assembly mechanisms and kinetic trapping [101], which some researchers take as a warning against trusting *in vitro* assembly results.

### 1.4.4 Unguided assembly models

While most physical assembly models fall into the previous two categories, there are a few researchers seeking to understand *how* viruses are able to assemble into the correct morphologies. A first step in this direction is to design geometries to target $T = 1$ assembly [102]. Because the angle differences between $T = 1$ and larger capsids are so large, no matching rules are required to achieve this easily. Elrad and Hagan [103] model $T = 1$ and $T = 3$ assembly around different-sized electrostatic cores. Their $T = 3$ results use two different conformations, but no matching rules. Moving to much larger capsids, we present in the following chapters our model [14] of irreversible assembly, which makes no assumption of final structure beyond preferring a certain angle between neighboring units. Levandovsky and Zandi [104] extended our model, primarily with improved growth rules, and were able to construct much larger capsids, including the characteristic cones of retroviruses.

One key aspect of unguided assembly models is that they often present a wide variety of failed assembly results: both in the form of capsids different from the assembly target (often because of being asymmetric), as well as unclosed "monsters" [14, 102]. Guided assembly models can experience failures as well [96, 105], but tend to retain the artificial network connectivity enforced by their rules (or else arise from allowing non-lattice-based interactions, as in [98]), and are therefore less interesting. Failures are particularly useful when comparing models to experiment. If two models both produce the same correct results, it may be the failures, compared with observed failure modes of actual viruses, that can best distinguish the models.

## 1.5 Atomistic approaches

The previous section represented a "top-down" approach, but as more protein structures are solved by X-ray diffraction and NMR, and as computational techniques and hardware improve, we have increasing opportunities to pursue a "bottom-up" approach, using atomistic techniques to determine the continuum behavior. In the following section we will review the relevant techniques, and then apply them in Chapter 6 to measure the elastic moduli of a sheet of capsid proteins, which we could then use in a continuum model. In Chapter 7 we go on to consider some of the subtleties of measuring diffusion and noise in the presence of hydrodynamic effects and thermostats.

All-atom MD has been applied to biomolecules for decades, ever since computers had sufficient memory to store the necessary data. The earliest MD biomolecule simulation was in 1977 on bovine pancreatic trypsin inhibitor [106], a 58aa protein whose structure was determined by X-ray crystallography. The simulation included four explicit water molecules (internal to the protein), while the rest of the water was im-

plicit, and ran for 9000 timesteps of approximately a femtosecond each. The primary conclusion was that the internals of a protein fluctuate diffusively, behaving in some ways more like a fluid than a solid, and that it is useful to think of a macromolecule with a ball-and-spring model. This software developed into the CHARMM (Chemistry at HARvard Macromolecular Mechanics) package [107], which has become a standard package for macromolecular simulations. Over the last three decades, increases in computation power have led to the all-atom MD simulation of an entire virus [108].

An opposite approach to MD is normal mode analysis (NMA). While the focus in MD is on integrating the equations of motion in time, the idea behind NMA is to evaluate the Hessian of the free energy, as a function of the atomic positions (this can also be coarse-grained to amino acid positions), about the equilibrium structure. This can be particularly useful in determining the principal modes of motion of a system (such as hinge axes of a protein) and their stiffnesses.

### 1.5.1 Force fields

Both MD and NMA make use of empirical potential energy functions. The most common potential used for proteins by many different packages is CHARMM22[12], which was calibrated from a wide variety of sources such as sampling of configurations in known structures, as well as *ab initio* techniques to determine the potential energy parameters [109]. The functional form of the potential involves electrostatics (with partial charges on each atom), along with a Lennard-Jones potential for long-range interactions. Bonded interactions have several terms, categorized by the number of bodies involved in each interaction. The two-body bond length interaction is simple

[12] the current version for nucleic acids is CHARMM27

harmonic; the three-body angle interaction is primarily harmonic, but is augmented in certain cases by a Urey-Bradley potential [110]; the four-body interactions include a cosine potential for dihedral angles, $V(\text{dihedral}) = k_\phi(1 + \cos(n\phi - \delta))$, although some dihedral angles are considered "improper" and involve a harmonic potential instead. The bonded parameters are specific to each different bond in each different residue, while the parameters in the Lennard-Jones potential are given by averaging the parameters of the two interacting atoms (geometrically for the strength, arithmetically for the length).

Other force fields in use are AMBER, GROMOS, and OPLS. Certain aspects of these force fields, such as the values of partial charges on atoms in amino acids, are converging in newer released parameters; Ponder and Case [111] suggest that simulations depending mostly on the repulsive interactions (such as our simulations of equilibrium fluctuations), should be roughly the same under any of the force fields, while simulations depending on more subtle features (e.g. ligand binding and protein folding simulations) may have stronger dependence on the force field.

One important aspect of the force field is whether it treats water explicitly or implicitly. Force fields for explicit water are designed with a particular water model in mind, and presumably lose accuracy when used with a different water model. The primary method of dealing with implicit solvent is the generalized Born model [112][13].

[13] Implicit water has a number of limitations, such as failing to correctly predict specific short-range effects, temperature and pressure, and (obviously) the effect of hydrodynamics and viscous drag. The CHARMM and AMBER packages both support implicit water, but as of this writing, NAMD has no plans to support it.

### 1.5.2 Molecular dynamics

Broadly, MD simulations are employed for a variety of uses, which we will categorize in terms of the importance of the dynamic aspect. First is the prediction of native structures (typically proteins, since most other macromolecules have well-known structures). This is most common (and successful) when starting from a crystal or NMR structure and then using an annealing approach to refine the structure to be more biologically relevant [113, 114], but also has farther-fetched applications in protein folding. In this use case, the dynamics are largely irrelevant: the end result is of utmost interest. In the opposite extreme are simulations focusing entirely on the dynamics. This category includes the techniques known as *steered* and *targeted* MD for exploring the pathways of rare transition events [115, 116, 113]. Finally, between these two extremes lie simulations of equilibrium ensembles, such as will be discussed later, in Chapters 6 and 7. In these simulations, we are interested in the distribution of structures in equilibrium, as well as the dynamics of the fluctuations.

We mentioned above two variations on simple MD, both of which involve adding "artificial" external forces to pull a system from one potential well into another. This is a common theme among MD variants. Another example of this occurs in multiscale models, where all-atom MD simulations are alternated with finite element or coarse-grained models at the boundaries. Villa et al. [117] explored protein–DNA systems by coarse-graining the DNA with a simple polymer model, recalculating the boundary forces every few MD timesteps.

Many of the examples listed above are more or less qualitative, in that the results consist in noticing certain types of motions, say a group of residues moving a certain direction, that connects the structure to the biological function. There have, however, been a number of quantitative results from molecular dynamics (particularly

from the steered simulations), most of which involve looking at mechanical measurements: *flexibility*, *stiffness*, and *elasticity*. One of the more interesting subjects is hinge-bending motions. Earlier work used only minimization techniques, holding the pertinent hinge at a fixed angle (and assuming the domains on either side to be rigid) to determine the bending potential in lysozyme [118], as well as in liver alcohol dehydrogenase [119]. Techniques like these for determining interdomain flexibilities are reviewed in the context of drug design by Carlson and McCammon [120].

One other technique is much more closely related to our work: essential dynamics (ED), also known as principal component analysis (PCA), uses equilibrium fluctuations similarly to us, diagonalizing the covariance matrix to find the largest-amplitude motions [121, 122, 123, 124], and disregarding the rest as unimportant. We both simplify and extend this approach, by assuming the important directions *a priori*, and by including the dynamics to give proper relative weight to each direction.

A number of researchers have critiqued ED on the grounds that it is manifestly dependent on the sampling window by demonstrating that the time series of their principal components look remarkably like cosines [125]. This is related to work on convergence in MD simulations and validation of trajectories [126, 127].

### 1.5.3 Normal mode analysis

NMA is the practice of starting with a potential energy function and computing a Hessian matrix which can then be diagonalized to find the normal modes and frequencies. While this has been a common technique in hard condensed matter since the 1950s [128], it was another several decades before biomolecules became accessible. The earlier normal mode analysis of a biomolecule was of the 29aa glucagon [129], using a specialized $\alpha$-helical force field with no non-bonded terms, and treating side

chains as point masses. Later studies took a variety of approaches: Go et al. [130] and Levitt et al. [131] studied 51aa of bovine pancreatic trypsin inhibitor (BPTI) by including all heavy atoms and polar hydrogens (needed for H-bonds) and using internal coordinates consisting of torsional and bond-length degrees of freedom, while Brooks and Karplus [132] treat the same protein using Cartesian degrees of freedom.

Once NMA was shown to work reasonably well with atomistic degrees of freedom and realistic potentials *in vacuo*, a number of simplifications and refinements were investigated.

One simplification by Tirion [133] is to use springs of a fixed stiffness between all pairs of atoms within a certain cutoff radius, rather than complicated empirical potentials[14], and has been shown to be roughly equivalent to ED [134][15]. This sort of approximation is sometimes known as an "elastic network model" or a "Gaussian network model", and provides a useful framework for calculating dynamics [135]. Later applications, including a few on viruses [136, 137] have taken the simplification even further, coarse-graining typically an entire residue into a single unit (this necessarily requires changing the cutoff lengths from the values used for atoms).

Under a Gaussian network model, Bahar et al. [138] show that vibration amplitude (and thus frequency) is related to wave number along the protein backbone. They observe that low-frequency fluctuation minima correspond to hinges (and deleterious mutations), while high-frequency maxima are the folding core.

This simplification is taken even further by treating the protein as a uniform volume of elastic material and then using finite element methods to understand the

[14]This makes it very clear that NMA depends heavily on having correct starting structures.

[15]While Rueda et al. [134] approached it differently, an interesting check of this uniform spring simplification would be to plot $\langle r^2 - \langle r \rangle^2 \rangle$ vs. $\langle r \rangle$ for all pairs of atoms in a protein.

elasticity [139, 140].

Another refinement is the inclusion of a damping term to the equations of motion, resulting in a solution to the Langevin dynamics. As such, Lamm and Szabo [141] worked out a formalism of "Langevin modes", in which they find a transition from underdamped to overdamped modes as a function of the viscosity. Kitao [142] extended this work by performing MD simulations in water and vacuum in addition to NMA, computing relaxation times similar to those we calculate in Chapter 6.

One important application of NMA is in building up coarse-grained models. Hayward et al. [143] constructed a general framework in which they break a system into blocks which are each approximated as a rigid body, described by a translation and a rotation, and then use the low-frequency modes to "dynamically" determine the best partitioning of the system. This was later extended by a clustering algorithm to allow automation of the technique [144]. Tama et al. [145] suggests that six residues per block is the limit to maintain the accuracy of the low-frequency modes. Essiz and Coalson [146] investigate the use of these block methods in Langevin dynamics.

## BIBLIOGRAPHY


[1] S. O. Kim, H. H. Solak, M. P. Stoykovich, N. J. Ferrier, J. J. de Pablo, and P. F. Nealey, "Epitaxial self-assembly of block copolymers on lithographically defined nanopatterned substrates", *Nature* **424**, 411 (2003).

[2] J. Y. Cheng, A. M. Mayes, and C. A. Ross, "Nanostructure engineering by templated self-assembly of block copolymers", *Nat. Mater.* **3**, 823 (2004).

[3] D. H. Gracias, J. Tien, T. L. Breen, C. Hsu, and G. M. Whitesides, "Forming Electrical Networks in Three Dimensions by Self-Assembly", *Science* **289**, 1170 (2000).

[4] G. M. Whitesides and B. Grzybowski, "Self-Assembly at All Scales", *Science* **295**, 2418 (2002).

[5] D. J. Kushner, "Self-assembly of biological structures", *Bacteriol. Rev.* **33**, 302 (1969).

[6] R. N. Perham, "Self-assembly of biological macromolecules", *Phil. Trans. R. Soc. B* **272**, 123 (1975).

[7] M. F. Carlier, "Actin polymerization and ATP hydrolysis", *Adv. Biophys.* **26**, 51 (1990).

[8] F. U. Hartl and J. Martin, "Molecular chaperones in cellular protein folding", *Curr. Opin. Struct. Biol.* **5**, 92 (1995).

[9] T. Ito, M. Bulger, M. J. Pazin, R. Kobayashi, and J. T. Kadonaga, "ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor", *Cell* **90**, 145 (1997).

[10] N. Kimizuka and T. Nakashima, "Spontaneous Self-Assembly of Glycolipid Bilayer Membranes in Sugar-philic Ionic Liquids and Formation of Ionogels", *Langmuir* **17**, 6759 (2001).

[11] J. N. Israelachvili, D. J. Mitchell, and B. W. Ninham, "Theory of self-assembly of lipid bilayers and vesicles", *Biochim. Biophys. Acta* **470**, 185 (1977).

[12] D. M. Salunke, D. L. D. Caspar, and R. L. Garcea, "Self-assembly of purified polyomavirus capsid protein VP1", *Cell* **46**, 895 (1986).

[13] R. F. Bruinsma, W. M. Gelbart, D. Reguera, J. Rudnick, and R. Zandi, "Viral Self-Assembly as a Thermodynamic Process", *Phys. Rev. Lett.* **90**, 248101 (2003).

[14] S. D. Hicks and C. L. Henley, "Irreversible growth model for virus capsid assembly", *Phys. Rev. E* **74**, 031912 (2006).

[15] S. D. Hicks and C. L. Henley, "Coarse-grained protein-protein stiffnesses and dynamics from all-atom simulations", *Phys. Rev. E* **81**, 030903 (2010).

[16] V. S. Reddy, H. A. Giesling, R. T. Morton, A. Kumar, C. B. Post, C. L. Brooks, and J. E. Johnson, "Energetics of Quasiequivalence: Computational Analysis of Protein-Protein Interactions in Icosahedral Viruses", *Biophys. J.* **74**, 546 (1998).

[17] F. H. C. Crick and J. D. Watson, "The Structure of Small Viruses", *Nature* **177**, 473 (1956).

[18] D. L. D. Caspar and A. Klug, "Physical Principles In Construction Of Regular Viruses", *Cold Spring Harb. Symp. Quant. Biol.* **27**, 1 (1962).

[19] S. Scarlata and C. Carter, "Role of HIV-1 Gag domains in viral assembly", *BBA-Biomembranes* **1614**, 62 (2003).

[20] E. O. Freed, "HIV-1 gag proteins: diverse functions in the virus life cycle", *Virology* **251**, 1 (1998).

[21] J. W. Wills and R. C. Craven, "Form, function, and use of retroviral gag proteins", *AIDS* **5**, 639 (1991).

[22] G. Cardone, J. G. Purdy, N. Cheng, R. C. Craven, and A. C. Steven, "Visualization of a missing link in retrovirus capsid assembly", *Nature* **457**, 694 (2009).

[23] J. G. Purdy, J. M. Flanagan, I. J. Ropson, K. E. Rennoll-Bankert, and R. C. Craven, "Critical role of conserved hydrophobic residues within the major homology region in mature retroviral capsid assembly", *J. Virol.* **82**, 5951 (2008).

[24] J. G. Purdy, J. M. Flanagan, I. J. Ropson, and R. C. Craven, "Retroviral Capsid Assembly: A Role for the CA Dimer in Initiation", *J. Mol. Biol.* **389**, 438 (2009).

[25] G. D. Bailey, J. K. Hyun, A. K. Mitra, and R. L. Kingston, "Proton-Linked Dimerization of a Retroviral Capsid Protein Initiates Capsid Assembly", *Structure* **17**, 737 (2009).

[26] S. Campbell and V. M. Vogt, "Self-assembly in vitro of purified CA-NC proteins from Rous sarcoma virus and human immunodeficiency virus type 1", *J. Virol.* **69**, 6487 (1995).

[27] B. K. Ganser, S. Li, V. Y. Klishko, J. T. Finch, and W. I. Sundquist, "Assembly and Analysis of Conical Models for the HIV-1 Core", *Science* **283**, 80 (1999).

[28] S. A. Datta, Z. Zhao, P. K. Clark, S. Tarasov, J. N. Alexandratos, S. J. Campbell, M. Kvaratskhelia, J. Lebowitz, and A. Rein, "Interactions between HIV-1 Gag Molecules in Solution: An Inositol Phosphate-mediated Switch.", *J. Mol. Biol.* **365**, 799 (2006).

[29] Y. M. Ma and V. M. Vogt, "Nucleic Acid Binding-Induced Gag Dimerization in the Assembly of Rous Sarcoma Virus Particles In Vitro", *J. Virol.* **78**, 52 (2004).

[30] C. T. Wang, H. Y. Lai, and J. J. Li, "Analysis of minimal human immunodeficiency virus type 1 gag coding sequences capable of virus-like particle assembly and release", *J. Virol.* **72**, 7950 (1998).

[31] M. C. Johnson, H. M. Scobie, Y. M. Ma, and V. M. Vogt, "Nucleic Acid-Independent Retrovirus Assembly Can Be Driven by Dimerization", *J. Virol.* **76**, 11177 (2002).

[32] P. W. Keller, M. C. Johnson, and V. M. Vogt, "Mutations in the Spacer Peptide and Adjoining Sequences in Rous Sarcoma Virus Gag Lead to Tubular Budding?", *J. Virol.* **82**, 6788 (2008).

[33] E. R. Wright, J. B. Schooler, H. J. Ding, C. Kieffer, C. Fillmore, W. I. Sundquist, and G. J. Jensen, "Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells.", *EMBO J.* **26**, 2218 (2007).

[34] D. Gheysen, E. Jacobs, F. D. Foresta, C. Thiriart, M. Francotte, D. Thines, and M. D. Wilde, "Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus-infected insect cells.", *Cell* **59**, 103 (1989).

[35] M. Ge and K. Sattler, "Observation of fullerene cones", *Chem. Phys. Lett.* **220**, 192 (1994).

[36] S. Li, C. P. Hill, W. I. Sundquist, and J. T. Finch, "Image reconstructions of helical assemblies of the HIV-1 CA protein.", *Nature* **407**, 409 (2000).

[37] J. Benjamin, B. K. Ganser-Pornillos, W. F. Tivol, W. I. Sundquist, and G. J. Jensen, "Three-dimensional structure of HIV-1 virus-like particles by electron cryotomography.", *J. Mol. Biol.* **346**, 577 (2005).

[38] J. A. Briggs, K. Grünewald, B. Glass, F. Förster, H. G. Kräusslich, and S. D. Fuller, "The Mechanism of HIV-1 Core Assembly: Insights from Three-Dimensional Reconstructions of Authentic Virions", *Structure* **14**, 15 (2006).

[39] J. A. G. Briggs, T. Wilk, R. Welker, H. G. Kräusslich, and S. D. Fuller, "Structural organization of authentic, mature HIV-1 virions and cores", *EMBO J.* **22**, 1707 (2003).

[40] C. Butan, D. C. Winkler, J. B. Heymann, R. C. Craven, and A. C. Steven, "RSV capsid polymorphism correlates with polymerization efficiency and envelope glycoprotein content: implications that nucleation controls morphogenesis", *J. Mol. Biol.* **376**, 1168 (2008).

[41] J. B. Heymann, C. Butan, D. C. Winkler, R. C. Craven, and A. C. Steven, "Irregular and semi-regular polyhedral models for Rous sarcoma virus cores", *Comput. Math. Methods Med.* **9**, 197 (2008).

[42] R. L. Kingston, N. H. Olson, and V. M. Vogt, "The organization of mature Rous sarcoma virus as studied by cryoelectron microscopy.", *J. Struct. Biol.* **136**, 67 (2001).

[43] B. K. Ganser-Pornillos, U. K. von Schwedler, K. M. Stray, C. Aiken, and W. I. Sundquist, "Assembly Properties of the Human Immunodeficiency Virus Type 1 CA Protein", *J. Virol.* **78**, 2545 (2004).

[44] I. Gross, H. Hohenberg, T. Wilk, K. Wiegers, M. Grattinger, B. Muller, S. Fuller, and H. G. Krausslich, "A conformational switch controlling HIV-1 morphogenesis.", *EMBO J.* **19**, 103 (2000).

[45] L. S. Ehrlich, T. Liu, S. Scarlata, B. Chu, and C. A. Carter, "HIV-1 Capsid Protein Forms Spherical (Immature-Like) and Tubular (Mature-Like) Particles in Vitro: Structure Switching by pH-induced Conformational Changes", *Biophys. J.* **81**, 586 (2001).

[46] J. A. G. Briggs, M. C. Johnson, M. N. Simon, S. D. Fuller, and V. M. Vogt, "Cryo-electron Microscopy Reveals Conserved and Divergent Features of Gag Packing in Immature Particles of Rous Sarcoma Virus and Human Immunodeficiency Virus", *J. Mol. Biol.* **355**, 157 (2006).

[47] D. Ako-Adjei, M. C. Johnson, and V. M. Vogt, "The Retroviral Capsid Domain Dictates Virion Size, Morphology, and Coassembly of Gag into Virus-Like Particles", *J. Virol.* **79**, 13463 (2005).

[48] B. K. Ganser-Pornillos, M. Yeager, and W. I. Sundquist, "The structural biology of HIV assembly", *Curr. Opin. Struct. Biol.* **18**, 203 (2008).

[49] O. Pornillos, B. K. Ganser-Pornillos, B. N. Kelly, Y. Hua, F. G. Whitby, C. D. Stout, W. I. Sundquist, C. P. Hill, and M. Yeager, "X-Ray Structures of the Hexameric Building Block of the HIV Capsid", *Cell* **137**, 1282 (2009).

[50] S. A. Datta, J. E. Curtis, W. Ratcliff, P. K. Clark, R. M. Crist, J. Lebowitz, S. Krueger, and A. Rein, "Conformation of the HIV-1 Gag Protein in Solution.", *J. Mol. Biol.* **365**, 812 (2007).

[51] E. Morita and W. I. Sundquist, "Retrovirus Budding", *Ann. Rev. Cell Dev. Biol.* **20**, 395 (2004).

[52] D. G. Demirov and E. O. Freed, "Retrovirus budding", *Virus Res.* **106**, 87 (2004).

[53] A. K. Dalton, P. S. Murray, D. Murray, and V. M. Vogt, "Biochemical Characterization of Rous Sarcoma Virus MA Protein Interaction with Membranes", *J. Virol.* **79**, 6227 (2005).

[54] A. K. Dalton, D. Ako-Adjei, P. S. Murray, D. Murray, and V. M. Vogt, "Electrostatic interactions drive membrane association of the HIV-1 Gag MA domain.", *J. Virol.* (2007).

[55] A. Alfadhli, D. Huseby, E. Kapit, D. Colman, and E. Barklis, "Human Immunodeficiency Virus Type 1 Matrix Protein Assembles on Membranes as a Hexamer?", *J. Virol.* **81**, 1472 (2007).

[56] D. R. Larson, M. C. Johnson, W. W. Webb, and V. M. Vogt, "Visualization of retrovirus budding with correlated light and electron microscopy", *Proc. Nat. Acad. Sci. USA* **102**, 15453 (2005).

[57] M. Gladnikoff and I. Rousso, "Directly Monitoring Individual Retrovirus Budding Events Using Atomic Force Microscopy", *Biophys. J.* **94**, 320 (2008).

[58] F. Bouamr, J. A. Melillo, M. Q. Wang, K. Nagashima, M. de Los Santos, A. Rein, and S. P. Goff, "PPPYEPTAP Motif Is the Late Domain of Human T-Cell Leukemia Virus Type 1 Gag and Mediates Its Functional Interaction with Cellular Proteins Nedd4 and Tsg101", *J. Virol.* **77**, 11882 (2003).

[59] M. C. Johnson, J. L. Spidel, D. Ako-Adjei, J. W. Wills, and V. M. Vogt, "The C-Terminal Half of TSG101 Blocks Rous Sarcoma Virus Budding and Sequesters Gag into Unique Nonendosomal Structures", *J. Virol.* **79**, 3775 (2005).

[60] C. Chen, O. Vincent, J. Jin, O. A. Weisz, and R. C. Montelaro, "Functions of Early (AP-2) and Late (AIP1/ALIX) Endocytic Proteins in Equine Infectious Anemia Virus Budding", *J. Biol. Chem.* **280**, 40474 (2005).

[61] I. L. Blanc, M. C. Prevost, M. C. Dokhelar, and A. R. Rosenberg, "The PPPY Motif of Human T-Cell Leukemia Virus Type 1 Gag Protein Is Required Early in the Budding Process", *J. Virol.* **76**, 10024 (2002).

[62] E. O. Freed, "Mechanisms of enveloped virus release.", *Virus Res.* **106**, 85 (2004).

[63] E. O. Freed, "Viral late domains", *J. Virol.* **76**, 4679 (2002).

[64] L. A. Carlson, J. A. G. Briggs, B. Glass, J. D. Riches, M. N. Simon, M. C. Johnson, B. Müller, K. Grünewald, and H. G. Kräusslich, "Three-Dimensional Analysis of Budding Sites and Released Virus Suggests a Revised Model for HIV-1 Morphogenesis", *Cell Host & Microbe* **4**, 592 (2008).

[65] E. G. Lee and M. L. Linial, "Deletion of a Cys-His motif from the Alpharetrovirus nucleocapsid domain reveals late domain mutant-like budding defects", *Virology* **347**, 226 (2006).

[66] C. Momany, L. C. Kovari, A. J. Prongay, W. Keller, R. K. Gitti, B. M. Lee, A. E. Gorbalenya, L. Tong, J. McClure, L. S. Ehrlich, et al., "Crystal structure of dimeric HIV-1 capsid protein", *Nat. Struct. Mol. Biol.* **3**, 763 (1996).

[67] T. R. Gamble, S. Yoo, F. F. Vajdos, U. K. von Schwedler, D. K. Worthylake, H. Wang, J. P. McCutcheon, W. I. Sundquist, and C. P. Hill, "Structure of the Carboxyl-Terminal Dimerization Domain of the HIV-1 Capsid Protein", *Science* **278**, 849 (1997).

[68] G. B. Mortuza, L. F. Haire, A. Stevens, S. J. Smerdon, and J. P. Stoye, "High-resolution structure of a retroviral capsid hexameric amino-terminal domain", *Nature* **431**, 481 (2004).

[69] M. Lu, S. C. Blacklow, and P. S. Kim, "A trimeric structural domain of the HIV-1 transmembrane glycoprotein", *Nat. Struct. Mol. Biol.* **2**, 1075 (1995).

[70] D. Ivanov, O. V. Tsodikov, J. Kasanov, T. Ellenberger, G. Wagner, and T. Collins, "Domain-swapped dimerization of the HIV-1 capsid C-terminal domain.", *Proc. Nat. Acad. Sci. USA* **104**, 4353 (2007).

[71] V. S. Reddy, H. A. Giesing, R. T. Morton, A. Kumar, C. B. Post, C. L. Brooks, and J. E. Johnson, "Energetics of Quasiequivalence: Computational Analysis of Protein-Protein Interactions in Icosahedral Viruses", *Biophys. J.* **74**, 546 (1998).

[72] V. M. Vogt and M. N. Simon, "Mass determination of rous sarcoma virus virions by scanning transmission electron microscopy.", *J. Virol.* **73**, 7050 (1999).

[73] F. Yu, S. M. Joshi, Y. M. Ma, R. L. Kingston, M. N. Simon, and V. M. Vogt, "Characterization of Rous Sarcoma Virus Gag Particles Assembled In Vitro", *J. Virol.* **75**, 2753 (2001).

[74] Y. Chen, B. Wu, K. Musier-Forsyth, L. M. Mansky, and J. D. Mueller, "Fluorescence fluctuation spectroscopy on viral-like particles reveals variable gag stoichiometry", *Biophys. J.* **96**, 1961 (2009).

[75] J. A. G. Briggs, M. N. Simon, I. Gross, H. G. Kraeusslich, S. D. Fuller, V. M. Vogt, and M. C. Johnson, "The stoichiometry of Gag protein in HIV-1", *Nat. Struct. Mol. Biol.* **11**, 672 (2004).

[76] J. Lanman, T. K. T. Lam, M. R. Emmett, A. G. Marshall, M. Sakalian, and P. E. P. Jr, "Key interactions in HIV-1 maturation identified by hydrogen-deuterium exchange", *J. Virol.* **78**, 2545 (2004).

[77] S. D. Fuller, T. Wilk, B. E. Gowen, H. G. Krausslich, and V. M. Vogt, "Cryo-electron microscopy reveals ordered domains in the immature HIV-1 particle", *Curr. Biol.* **7**, 729 (1997).

[78] J. A. G. Briggs, J. D. Riches, B. Glass, V. Bartonova, G. Zanetti, and H. G. Kräusslich, "Structure and assembly of immature HIV", *Proc. Nat. Acad. Sci. USA* **106**, 11090 (2009).

[79] M. M. Inamdar, W. M. Gelbart, and R. Phillips, "Dynamics of DNA Ejection from Bacteriophage", *Biophys. J.* **91**, 411 (2006).

[80] Y. R. Chemla, K. Aathavan, J. Michaelis, S. Grimes, P. J. Jardine, D. L. Anderson, and C. Bustamante, "Mechanism of force generation of a viral DNA packaging motor", *Cell* **122**, 683 (2005).

[81] S. Tzlil, M. Deserno, W. M. Gelbart, and A. Ben-Shaul, "A Statistical-Thermodynamic Model of Viral Budding", *Biophys. J.* **86**, 2037 (2004).

[82] L. Pauling, "Protein interactions. Aggregation of globular proteins", *Disc. Faraday Soc.* **13**, 170 (1953).

[83] D. Cavanagh, "Coronaviridae: a review of coronaviruses and toroviruses", *Coronaviruses with Special Emphasis on First Insights Concerning SARS* pp. 1–54 (2005).

[84] J. E. Johnson and J. A. Speir, "Quasi-equivalent viruses: a paradigm for protein assemblies", *J. Mol. Biol.* **269**, 665 (1997).

[85] R. Twarock, "A tiling approach to virus capsid assembly explaining a structural puzzle in virology", *J. Theor. Biol.* **226**, 477 (2004).

[86] R. Zandi, D. Reguera, R. F. Bruinsma, W. M. Gelbart, and J. Rudnick, "Origin of icosahedral symmetry in viruses", *Proc. Nat. Acad. Sci. USA* **101**, 15556 (2004).

[87] T. T. Nguyen, R. F. Bruinsma, and W. M. Gelbart, "Elasticity theory and shape transitions of viral shells", *Phys. Rev. E* **72**, 051923 (2005).

[88] T. T. Nguyen, R. F. Bruinsma, and W. M. Gelbart, "Continuum Theory of Retroviral Capsids", *Phys. Rev. Lett.* **96**, 078102 (2006).

[89] J. Lidmar, L. Mirny, and D. R. Nelson, "Virus shapes and buckling transitions in spherical shells", *Phys. Rev. E* **68**, 051910 (2003).

[90] A. Zlotnick, "To build a virus capsid. An equilibrium model of the self assembly of polyhedral protein complexes.", *J. Mol. Biol.* **241**, 59 (1994).

[91] A. Zlotnick, "Theoretical aspects of virus capsid assembly", *J. Mol. Recog.* **18**, 479 (2005).

[92] D. Endres and A. Zlotnick, "Model-Based Analysis of Assembly Kinetics for Virus Capsids or Other Spherical Polymers", *Biophys. J.* **83**, 1217 (2002).

[93] D. Endres, M. Miyahara, P. Moisant, and A. Zlotnick, "A reaction landscape identifies the intermediates critical for self-assembly of virus capsids and other polyhedral structures", *Protein Sci.* **14**, 1518 (2005).

[94] A. Morozov, R. Bruinsma, and J. Rudnick, "Assembly of Viruses and the Pseudo Law of Mass Action", *Biophys. J.* **96**, 419 (2009).

[95] T. Keef, A. Taormina, and R. Twarock, "Assembly models for Papovaviridae based on tiling theory", *Phys. Biol.* **2**, 175 (2005).

[96] B. Berger, P. W. Shor, L. Tucker-Kellogg, and J. King, "Local Rule-Based Theory of Virus Shell Assembly", *Proc. Nat. Acad. Sci. USA* **91**, 7732 (1994).

[97] R. Schwartz, P. W. Shor, P. E. Prevelige, and B. Berger, "Local Rules Simulation of the Kinetics of Virus Capsid Self-Assembly", *Biophys. J.* **75**, 2626 (1998).

[98] H. D. Nguyen and C. L. B. III, "Generalized Structural Polymorphism in Self-Assembled Viral Particles", *Nano Lett.* **8**, 4574 (2008).

[99] R. Schwartz, P. W. Shor, and B. Berger, "Local rule simulations of capsid assembly", *J. Theor. Med.* **6**, 81 (2005).

[100] D. C. Rapaport, "Self-assembly of polyhedral shells: A molecular dynamics study", *Phys. Rev. E* **70**, 51905 (2004).

[101] B. Sweeney, T. Zhang, and R. Schwartz, "Exploring the Parameter Space of Complex Self-Assembly through Virus Capsid Models", *Biophys. J.* **94**, 772 (2008).

[102] M. F. Hagan and D. Chandler, "Dynamic Pathways for Viral Capsid Assembly", *Biophys. J.* **91**, 42 (2006).

[103] O. M. Elrad and M. F. Hagan, "Mechanisms of Size Control and Polymorphism in Viral Capsid Assembly", *Nano Lett.* **8**, 3850 (2008).

[104] A. Levandovsky and R. Zandi, "Nonequilibirum Assembly, Retroviruses, and Conical Structures", *Phys. Rev. Lett.* **102**, 198102 (2009).

[105] H. D. Nguyen, V. S. Reddy, and C. L. B. III, "Invariant polymorphism in virus capsid assembly", *J. Am. Chem. Soc.* **131**, 2606 (2009).

[106] J. A. McCammon, B. R. Gelin, and M. Karplus, "Dynamics of folded proteins", *Nature* **267**, 585 (1977).

[107] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, "CHARMM: a program for macromolecular energy, minimization, and dynamics calculations", *J. Comput. Chem.* **4**, 187 (1983).

[108] P. L. Freddolino, A. S. Arkhipov, S. B. Larson, A. McPherson, and K. Schulten, "Molecular Dynamics Simulations of the Complete Satellite Tobacco Mosaic Virus", *Structure* **14**, 437 (2006).

[109] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, et al., "All-atom empirical potential for molecular modeling and dynamics studies of proteins", *J. Phys. Chem. B* **102**, 3586 (1998).

[110] H. C. Urey and C. A. Bradley, "The vibrations of pentatonic tetrahedral molecules", *Phys. Rev.* **38**, 1969 (1931).

[111] J. W. Ponder and D. A. Case, "Force fields for protein simulations.", *Adv. Protein Chem.* **66**, 27 (2003).

[112] D. Bashford and D. A. Case, "Generalized Born models of macromolecular solvation effects", *Ann. Rev. Phys. Chem.* **51**, 129 (2000).

[113] M. Karplus and J. A. McCammon, "Molecular dynamics simulations of biomolecules", *Nat. Struct. Biol.* **9**, 646 (2002).

[114] A. T. Brunger, J. Kuriyan, and M. Karplus, "Crystallographic R Factor Refinement by Molecular Dynamics.", *Science* **235**, 458 (1987).

[115] J. Schlitter, M. Engels, and P. Kruger, "Targeted molecular dynamics: a new approach for searching pathways of conformational transitions.", *J. Mol. Graph.* **12**, 84 (1994).

[116] B. Isralewitz, M. Gao, and K. Schulten, "Steered molecular dynamics and mechanical functions of proteins", *Curr. Opin. Struct. Biol.* **11**, 224 (2001).

[117] E. Villa, A. Balaeff, L. Mahadevan, and K. Schulten, "Multiscale method for simulating protein-DNA complexes", *Multiscale Model. Simul.* **2**, 527 (2004).

[118] J. A. McCammon, B. R. Gelin, M. Karplus, and P. G. Wolynes, "The hinge-bending mode in lysozyme", *Nature* **262**, 325 (1976).

[119] F. Colonna-Cesari, D. Perahia, M. Karplus, H. Eklund, C. I. Braden, and O. Tapia, "Interdomain motion in liver alcohol dehydrogenase. Structural and energetic analysis of the hinge bending mode", *J. Biol. Chem.* **261**, 15273 (1986).

[120] H. A. Carlson and J. A. McCammon, "Accommodating Protein Flexibility in Computational Drug Design", *Mol. Pharmacol.* **57**, 213 (2000).

[121] A. Amadei, A. B. Linssen, and H. J. Berendsen, "Essential dynamics of proteins.", *Proteins* **17**, 412 (1993).

[122] T. Horiuchi and N. Go, "Projection of Monte Carlo and molecular dynamics trajectories onto the normal mode axes: human lysozyme", *Proteins* **10**, 106 (1991).

[123] T. Ichiye and M. Karplus, "Collective motions in proteins: A covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations", *Proteins Str. Fun. Gen.* **11**, 205 (1991).

[124] A. Kitao and N. Go, "Investigating protein dynamics in collective coordinate space", *Curr. Opin. Struct. Biol.* **9**, 164 (1999).

[125] M. A. Balsera, W. Wriggers, Y. Oono, and K. Schulten, "Principal Component Analysis and Long Time Protein Dynamics", *J. Phys. Chem.* **100**, 2567 (1996).

[126] B. Hess, "Convergence of sampling in protein simulations", *Phys. Rev. E* **65**, 031910 (2002).

[127] B. Hess, "Similarities between principal components of protein dynamics and random diffusion", *Phys. Rev. E* **62**, 8438 (2000).

[128] E. B. Wilson, J. C. Decius, and P. C. Cross, *Molecular vibrations* (McGraw-Hill, New York, 1955).

[129] M. Tasumi, H. Takeuchi, S. Ataka, A. M. Dwivedi, and S. Krimm, "Normal vibrations of proteins: glucagon.", *Biopolymers* **21**, 711 (1982).

[130] N. Go, T. Noguti, and T. Nishikawa, "Dynamics of low-frequency vibrational modes", *Proc. Nat. Acad. Sci. USA* **80**, 3696 (1983).

[131] M. Levitt, C. Sander, and P. S. Stern, "The normal modes of a protein: Native bovine pancreatic trypsin inhibitor", *Int. J. Quant. Chem.* **10**, 181 (1983).

[132] B. Brooks and M. Karplus, "Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor", *Proc. Nat. Acad. Sci. USA* **80**, 6571 (1983).

[133] M. M. Tirion, "Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis", *Phys. Rev. Lett.* **77**, 1905 (1996).

[134] M. Rueda, P. Chacón, and M. Orozco, "Thorough Validation of Protein Normal Mode Analysis: A Comparative Study with Essential Dynamics", *Structure* **15**, 565 (2007).

[135] T. Haliloglu, I. Bahar, and B. Erman, "Gaussian Dynamics of Folded Proteins", *Phys. Rev. Lett.* **79**, 3090 (1997).

[136] A. J. Rader, D. H. Vlad, and I. Bahar, "Maturation Dynamics of Bacteriophage HK97 Capsid", *Structure* **13**, 413 (2005).

[137] F. Tama and C. L. Brooks, "The Mechanism and Pathway of pH Induced Swelling in Cowpea Chlorotic Mottle Virus", *J. Mol. Biol.* **318**, 733 (2002).

[138] I. Bahar, A. R. Atilgan, M. C. Demirel, and B. Erman, "Vibrational Dynamics of Folded Proteins: Significance of Slow and Fast Motions in Relation to Function and Stability", *Phys. Rev. Lett.* **80**, 2733 (1998).

[139] M. M. Gibbons and W. S. Klug, "Mechanical modeling of viral capsids", *J. Mat. Sci.* **42**, 8995 (2007).

[140] M. M. Gibbons and W. S. Klug, "Nonlinear finite-element analysis of nanoindentation of viral capsids", *Phys. Rev. E* **75**, 031901 (2007).

[141] G. Lamm and A. Szabo, "Langevin modes of macromolecules", *J. Chem. Phys.* **85**, 7334 (1986).

[142] A. Kitao, "The effects of solvent on the conformation and the collective motions of protein: normal mode analysis and molecular dynamics simulations of melittin in water and in vacuum", *Chem. Phys.* **158**, 447 (1991).

[143] S. Hayward, A. Kitao, and H. J. C. Berendsen, "Model-Free Methods of Analyzing Domain Motions in Proteins From Simulation: A Comparison of Normal Mode Analysis and Molecular Dynamics Simulation of Lysozyme", *Proteins Str. Fun. Gen.* **27**, 425 (1997).

[144] S. Hayward and H. J. C. Berendsen, "Systematic Analysis of Domain Motions in Proteins From Conformational Change: New Results on Citrate Synthase and T4 Lysozyme", *Proteins* **27**, 425 (1997).

[145] F. Tama, F. X. Gadea, O. Marques, and Y. H. Sanejouand, "Building-block approach for determining low-frequency normal modes of macromolecules", *Proteins Str. Fun. Gen.* **41**, 1 (2000).

[146] S. G. Essiz and R. D. Coalson, "Langevin dynamics of molecules with internal rigid fragments in the harmonic regime", *J. Chem. Phys.* **127**, 104109 (2007).

CHAPTER 2

# AN IRREVERSIBLE GROWTH MODEL OF CAPSID ASSEMBLY

This chapter consists of much of our original assembly paper [1], minus the introduction, which has been integrated into the previous chapter. The only addition is Figure 2.4.

## 2.1 Irreversible Growth Model

We now introduce a model to describe quasiequivalent capsid assembly in a far-from-equilibrium picture. Consider a single growing capsid and a number of units in solution. We picture the units slowly accreting onto the growing capsid until the finished product is formed.

Our choice is to represent this by adopting the simplest possible model that can represent a growing capsid and be simulated efficiently: this precludes representing each protein as a rigid body moving in space. Instead, a capsid (growing or completed) is represented by a triangular network (§2.1.1), with an elastic energy governing the bond lengths and angles (§2.1.2). We do not explicitly represent the units in solution, instead formulating a set of first-order rate equations for addition of a unit to the capsid or for other discrete changes in the network geometry (§2.1.3).

Other physical or mathematical models have been abstracted to a similar degree [2, 3, 4], following a standard philosophy of statistical mechanics. Some capacity to adapt the model to (say) a specific virus species is lost, but the simplicity makes it conceivable to grasp the physical meaning of each parameter, and feasible to explore all dimensions of the parameter space by simulations. Typically, only particular combinations of the microscopic parameters matter, and a properly formulated toy model adopts those combinations. It can happen that fairly different microscopic

systems may, through such elimination of unimportant parameters, all map to the same simple model; in that case, the model offers a possibility of unifying the description of all these systems.

### 2.1.1 Configuration degrees of freedom

Our formulation depends on *two* complementary kinds of degree of freedom, a discrete kind we call "topological" and a continuous kind called "positional". The former consists of a bond network built from triangles, with vertices either connected by a bond or not; the latter consists of the actual coordinates of the vertices in space. Since prior work emphasized equilibrium, we took the opposite limit by allowing no change in any bond, once formed. One consequence is that our discrete "topological" variables are more fundamental than the positional ones: given a network of bonds, the angles and bond lengths will relax to a constrained minimum (or fluctuate thermally around it) as determined by a Hamiltonian, defined in §2.1.2. In our growth model, these positional variables feed back into the discrete ones by controlling the relative rates of alternative changes in the network as units are accreted. (In principle one could envisage a further abstraction in which the positional variables are eliminated completely and the rates expressed directly in terms of the bond topology, but we did not attempt that.)

The models discussed above in §1.4 all have essentially just a single type of degree of freedom—positional, topological, or something intermediate. Lidmar et al. [4] assume a predetermined graph topology, so only the vertices' positions are nontrivial. On the other hand, Endres et al. [5] discard position information and only consider the (discrete) connectivity. Finally, Bruinsma et al. [2] continuously vary the positions of the discs, and determine which discs neighbor one another secondarily.

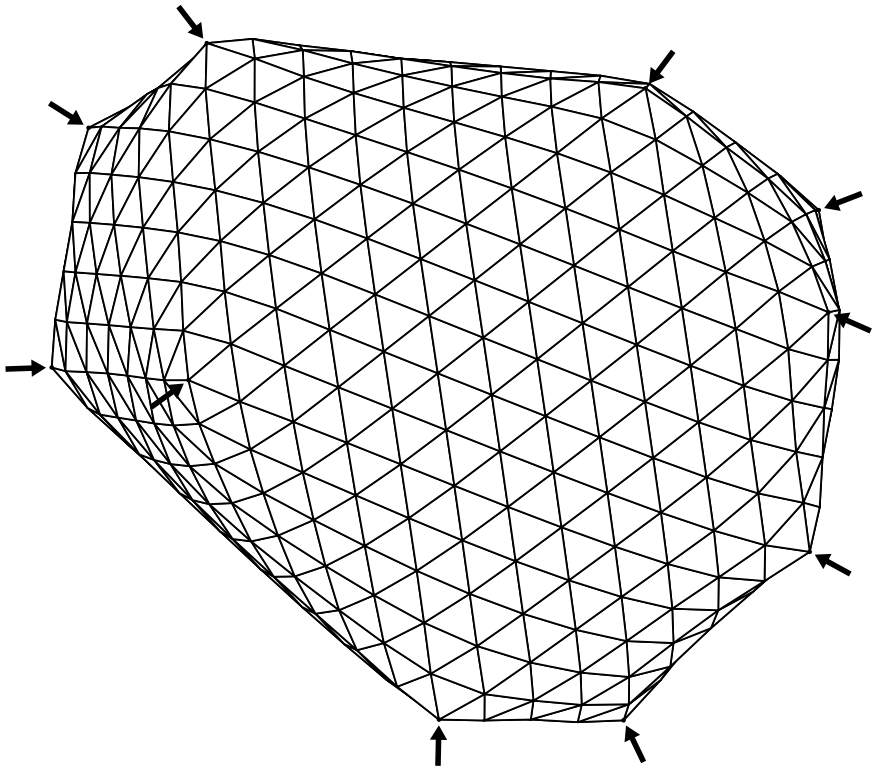


Figure 2.1: Example of a closed final capsid resulting from our growth simulation, with $\ell_f = .12 r_0$, $\theta_0 = 16°$, $\Gamma_{I,J} = 50$, and $\sigma = 12.7°$. These parameters are explained in §§2.1.2 and 2.1.3. The disclinations are marked by arrows.

It may be questioned why we have chosen triangles as the fundamental building blocks. In a model more faithful to a particular virus species, one would want to add the multimer that is accreted in nature. Virus species assembling from dimers [6, 7], trimers [8], and pentamers/hexamers [9], have all been observed experimentally. HIV has so far formed only dimers in solution [10]. Several groups have done molecular dynamics simulations using solutions of monomers [11, 12] and kinetic simulations with pentamers [13], dimers [14], or trimers [15]. Because of its simplicity, and the work done on tethered surfaces by Nelson and coworkers [16, 4], we will focus on a trimer-based model for this initial work, an example of which can be seen in Figure 2.1. It is at this point worth noting that the hexagonal lattice that is dual to

our triangular lattice is in fact very similar to the molecular lattices formed from HIV CA [17].

Our use of triangle units is also influenced by the notion of "universality" in physics, whereby the functional form of elastic theory, or the critical exponents of a phase transition, are independent of the particular lattice used at the microscopic scale. In any of the alternative representations, one can still define a triangular, locally sixfold lattice with rare locally 5-fold points in it. Much experience in statistical mechanics suggests that, at "coarse-grained" length scales (those large compared to the lattice spacing), the behavior stops depending on the details. However, two related caveats must be expressed, that (i) possibly a detail of the microscopic model forces a certain parameter of the coarse-grained model to be strictly zero, thereby changing the qualitative behavior ("universality class"); (ii) it may be that a parameter regime easy to achieve in one version of the model will require a complicated fine-tuning of parameters in an alternate version.

Caspar and Klug [18] noted that only pentamers and hexamers have small enough deformations to be allowed in quasiequivalence, and thus any quasiequivalent capsid must have exactly twelve pentamers. Quasiequivalence is based on a flat triangular lattice, so that a pentamer is a *disclination*: a topological defect of the triangular lattice. This means it can be characterized by effects at an arbitrary distance; namely, if we parallel transport a vector around a loop, that vector ends up rotated by $(\pi/3)N_{\mathrm{disc}}$ from its starting orientation, where $N_{\mathrm{disc}}$ is the number of 5-fold disclinations enclosed by the loop. (This is called a "disclination charge" by analogy to how the electric charge enclosed by a surface is determined by an integral of the electric field over that surface, according to Gauss's law.) That there are exactly twelve disclinations can now be seen either by counting vertices, edges, and triangles under the

constraint $V - E + F = 2$, or more generally because the total disclination charge must sum to $4\pi$.[1]

Between the disclinations are patches of regular sixfold lattice with no topological freedom: thus, *the capsid is completely determined by the placement of the disclinations*. Since there may be hundreds of network vertices, and only twelve disclinations, this is in principle a simplification.

### 2.1.2 Hamiltonian

We represent the growing capsid as a number of approximately equilateral triangles connected along the edges. We then generalize the discretized Hamiltonian used by Lidmar et al. [4] to include spontaneous curvature $\theta_0$ and steric terms.

$$\mathscr{H} = \mathscr{H}_{\text{stretch}} + \mathscr{H}_{\text{bend}} + \mathscr{H}_{\text{steric}}. \tag{2.1}$$

#### 2.1.2.1 Elastic energy

The first two terms in (2.1) are elastic terms for bond stretching and bending:

$$\mathscr{H}_{\text{stretch}} = \frac{\sqrt{3}\widetilde{Y}}{4} \sum_{\langle ij \rangle} \left( \left| \boldsymbol{r}_i - \boldsymbol{r}_j \right| - r_0 \right)^2, \tag{2.2}$$

$$\mathscr{H}_{\text{bend}} = \frac{2\widetilde{\kappa}}{\sqrt{3}} \sum_{\langle IJ \rangle} \left( 1 - \cos\left( \theta_{IJ} - \theta_0 \right) \right). \tag{2.3}$$

Here, $\langle ij \rangle$ denote pairs of nearest-neighbor vertices with positions $\boldsymbol{r}_i$, and $\langle IJ \rangle$ denote pairs of nearest-neighbor triangles. The exterior dihedral angle

$$\theta_{IJ} = \cos^{-1}\left( \hat{\boldsymbol{n}}_I \cdot \hat{\boldsymbol{n}}_J \right), \tag{2.4}$$

[1] These are both consequences of the Euler characteristic for a genus 0 surface, $\chi(g) = 2 - 2g = 2$. More generally, $V - E + F = \chi(g)$, and the total disclination charge must sum to $2\pi\chi(g)$. For more information, see Eric W. Weisstein. "Euler Characteristic." From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/EulerCharacteristic.html and references therein

where $\hat{\boldsymbol{n}}_I$ is the unit normal to triangle $I$.

Our discretized parameters, $\widetilde{Y}$ and $\widetilde{\kappa}$, have the same dimensions as the two-dimensional Young's modulus $Y$, and bending stiffness $\kappa$, respectively, which are emphasized in continuum approaches to predicting capsid shapes [4], and for a flat sheet in the linearized regime the parameters have the same values as well. If we parameterized our model by spring constants $K_{\text{stretch}}$ and $K_{\text{bend}}$ equal to the curvature of our radial and angular potentials at the bottoms of their respective wells, we would have $\widetilde{Y} = \frac{2}{\sqrt{3}} K_{\text{stretch}}$ and $\widetilde{\kappa} = \frac{\sqrt{3}}{2} K_{\text{bend}}$. In most cases (except in direct comparison with some experimental measurements), we are only concerned with the ratio of these elastic constants. This ratio provides a length scale, the *Foppl-von Kármán length*,

$$\ell_{\text{f}}^2 \equiv \kappa / Y. \tag{2.5}$$

From this point on, we will take units such that $r_0 = 1$, and therefore the parameter $\ell_{\text{f}}$ is effectively dimensionless.

Previous work made this same ratio dimensionless using the capsid *radius* $R$, rather than the triangle size $r_0$, and thus defined the Foppl-von Kármán number [4]

$$\gamma = Y R^2 / \kappa. \tag{2.6}$$

The capsid radius $R$ is well-defined in the case of a spherical capsid, but for non-spherical capsids, a definition of $R$ is problematic; and in any case, $\ell_{\text{f}}$ controls many other properties, such as the exponential decay of strain and Gaussian curvature with distance from a defect. Thus, we consider $\ell_{\text{f}}$ to be the more fundamental parameter and thus write $\gamma = (R/\ell_{\text{f}})^2$. We note that a small $\ell_{\text{f}}$ corresponds to a large Young's modulus and therefore an *angular* (or faceted) regime. On the other hand, large $\ell_{\text{f}}$ entails a large bending stiffness and leads to a *smooth* regime [4]. Since our model

is two-dimensional, we are able to specify arbitrarily large $\ell_{\mathrm{f}}$. Physically, however, $\ell_{\mathrm{f}}$ must be on the order of the capsid thickness or smaller.

#### 2.1.2.2 Spontaneous curvature and steric repulsion

Microscopically, we expect that capsid proteins are more similar in shape to cones or pyramids, with the apex toward the inside, than to cylinders[2]. Therefore, if two proteins are in contact, the outer surface will be bent at a characteristic angle. This suggests that $\mathscr{H}_{\mathrm{bend}}$ should favor some dihedral angle $\theta_0$, appearing in (2.3). Additionally, it motivates our model of steric repulsion based on tetrahedra, explained below.

The preferred dihedral angle $\theta_0$ is a key parameter since it is the main determinant of capsid size in our model, as was speculated to be the case in real capsids [18]. This corresponds to spontaneous curvature in a continuum model.

The final term $\mathscr{H}_{\mathrm{steric}}$ in (2.1) is a steric potential, chosen to vanish for all physically realistic capsids. The steric potential proves difficult to incorporate into our cartoon model, for two reasons. Firstly, all the inter-unit interactions of properly bonded units should already be accounted for in the elastic term $\mathscr{H}_{\mathrm{stretch}} + \mathscr{H}_{\mathrm{bend}}$, so we demand that the steric force not make additional contributions to these forces. Secondly, the other terms in (2.1) relate units that are "topological neighbors", as defined by the bond network (the discrete configuration). But two parts of the capsid that are distant topologically may grow to be nearby in real space (the positional configuration), and must then be kept from intersecting. Thus, the steric term must apply equally to topologically distant segments of the capsid, or to adjacent units, e.g. two

[2]Indeed, the CA protein in HIV consists of two separate domains, arranged in a wedge shape with the smaller C–terminal domain pointing towards the inside of the capsid and the larger N–terminal domain on the outside. See, for instance, the structures in [10]. RSV is similar [19]

as-yet unjoined triangles on the same vertex.

To implement a computationally tractable steric potential, we imagine each triangle to be the base of an inward-pointing tetrahedron, and add a repulsion between the apex of each tetrahedron and the vertices on the base of each other. This potential vanishes for physically realistic capsids. A more technical discussion may be found in §2.6, and the steric Hamiltonian is defined in (2.23).

#### 2.1.2.3 Microscopic estimation of elastic energy

Interactions between capsid proteins have been simulated electrostatically [20] to determine binding energies for large multimers of capsid proteins, necessarily in different relative positions. Such simulations could be extended to determine the elastic constants for particular viruses with known protein structure.

Alternatively, we can perform a rough estimate of the elastic parameters by considering some experimental measurements. Vliegenthart and Gompper [21] performed extensive computational studies to relate experiments with atomic force microscope (AFM) indentation of capsids to a model very similar to ours. Thus, we can use these AFM studies to determine the appropriate magnitude of $Y$ and $\kappa$. Ivanovska et al. [22] carried out mechanical structure measurements on the $T = 3$ phage $\phi 29$ and found the bulk modulus $B \approx 1.4$GPa and the thickness $t \approx 2.5$nm. We obtain an estimate of the two-dimensional Young's modulus by $Y \sim Bt \approx 3.5$N/m [23].

We could also estimate the elastic parameters from persistence length measurements. Maeda and Fujime [24] measured the tube-forming phage *fd* in suspension and determined the persistence length of the 9nm-diameter tubes at 22°C to be 3.9μm. If we construct a tube out of our triangular units, the persistence length would

be

$$\xi_p \approx \frac{R}{k_B T}\left(\kappa + \frac{8}{9} Y R^2\right), \tag{2.7}$$

where $R$ is the radius of the tube. Thus, we can conclude

$$\kappa + (8/9) Y R^2 = Y\left((8/9)R^2 + \ell_f^2\right) \approx 22\text{eV}, \tag{2.8}$$

which puts an upper bound of 0.17N/m on $Y$, in sharp contrast to the $\phi 29$ results above. Moreover, since *fd* is charged [25], the purely elastic contribution to the persistence length may be much smaller, making our estimate very conservative. If we previously knew either $\ell_f$ or one of $Y$ or $\kappa$, we could use this measurement to determine the others.

To get an idea of the elastic parameters for HIV, we can produce model capsids by hand that resemble HIV cores. In particular, we grew several capsids with about 500 triangles in a cone shape. Tuning the elastic parameters to roughly match the observed shape of HIV [26, 21], we found $\gamma = (R/\ell_f)^2 \approx 550$ produces the correct amount of facetedness. This corresponds to $r_0/\ell_f = 6$. Using our results for capsid size as a function of $\ell_f$ and $\theta_0$, presented in §2.3.1, we can guess that such a capsid would require $\theta_0 \approx 20°$ to be grown by our model.

Given a set of connected triangles (a topological configuration), we can now use this Hamiltonian to determine the lowest-energy configuration of the positions of the triangles. These positions correspond to a continuous degree of freedom that is now fully determined by the model ($\mathscr{H}$) and the connectivities—the discrete degree of freedom. Ultimately, we are only concerned with the discrete configuration.

### 2.1.3 Growth

We have noted that capsids are determined by the locations of the disclinations (pentamers). For an irreversible growth model, in which no step can be undone, the fundamental question is therefore: while growth occurs at the border, which twelve vertices are frozen in as pentamers? Keeping this in mind, we will now discuss our capsid growth process.

#### 2.1.3.1 Growth steps

We define transitions between incomplete capsids, consistent with irreversible growth, called *growth steps*. Two elementary growth steps are immediately apparent: *accretion* and *joining*. Accretion is the addition of a single triangle to a border edge and joining is the formation of a bond between two adjacent border edges. We require the vertex between these two joined edges to have five or six triangles around it in order to ensure that only pentamers and hexamers form.

Besides accretion and joining, we define a third, composite growth step: *insertion*. We define insertion as accretion followed by joining along an edge of the new triangle. The vertex into which we insert must have four or five triangles. Insertion at a 4-vertex or joining at a 5-vertex is the only way to form a pentamer. These three steps are illustrated in Figure 2.2.

Growth begins with a small template—either a single triangle or a pentamer of five triangles about a vertex. From here, the growth is determined by the sequence of growth steps, which is chosen stochastically. We will first present our rules for the relative probability of choosing the growth steps, and then explain their microscopic rationalization.

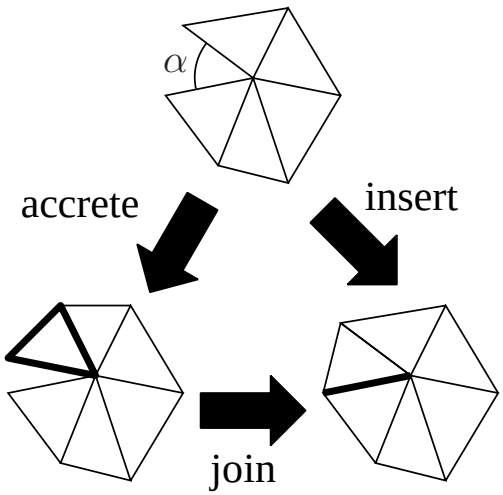



Figure 2.2: Elementary growth steps of *insertion*, *joining*, and *accretion*, performed from the same starting point of an edge with opening angle $\alpha$. Insertion can be decomposed into Accretion followed by Joining.

#### 2.1.3.2 Rates

We precede each growth step by relaxing all vertex positions using a conjugate-gradient algorithm to minimize the positional energy $\mathscr{H}$. Now a rate $k_\nu$ is defined for each allowed growth step $\nu$, which is a function of the local topology and of the *opening angle* $\alpha$ between pairs of edges at each vertex on the border, defined in Figure 2.2. The probability of step $\nu$ is then taken to be $k_\nu / \sum_\mu k_\mu$; once a step $\nu$ is picked, we perform the step and iterate the process (beginning as before with a relaxation).

We take the accretion rate, $k_\mathrm{A}$, to be independent of the local configuration: in particular it is not a function of $\alpha$.

So long as we are concerned only with the outcome and not the time taken to reach it, only relative rates are relevant. Thus we can now define

$$\frac{k_\mathrm{J}(\alpha)}{k_\mathrm{A}} = \Gamma_\mathrm{J} \mathrm{e}^{-\alpha^2/2\sigma^2} \tag{2.9}$$

$$\frac{k_\mathrm{I}(\alpha)}{k_\mathrm{A}} = \Gamma_\mathrm{I} \mathrm{e}^{-(\alpha-\pi/3)^2/2\sigma^2}, \tag{2.10}$$

with justification to follow. Note that steps are only considered if (1) they do not break

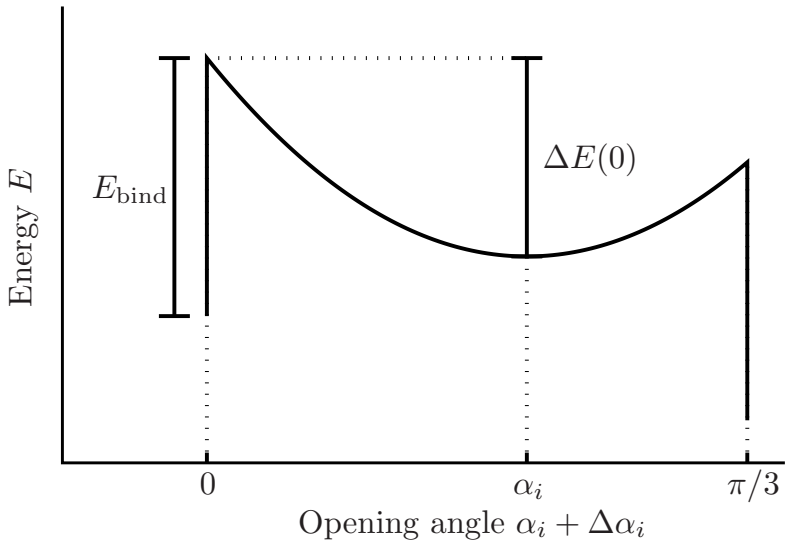



Figure 2.3: A representation of the energy as a function of opening angle fluctuations $\Delta\alpha_i$. When $\alpha_i + \Delta\alpha_i$ reaches either 0 or $\pi/3$, we imagine joining or insertion, respectively, occurring, reducing the total energy by $E_{\text{bind}}$, for joining, and some combination of $E_{\text{bind}}$ and a chemical potential for insertion.

any topological rules by enclosing a non-pentamer/hexamer, and (2) they do not lead to steric hindrances. This second point is discussed further in §2.6.

#### 2.1.3.3 Microscopic justification of rates

While many models explicitly account for units in solution and fluctuations in incomplete capsids [27, 12], we have chosen a simplified cartoon model. Implicit to this is the idea that the capsid is thermally fluctuating between growth steps.

Say the time between successive additions is longer than the relaxation time scale of the positional degrees of freedom. Then between each growth step, we can assume that the incomplete capsid is in equilibrium and thus samples a Boltzmann distribution. We consider the energy of fluctuations about the relaxed position. For a particular vertex, $i$, with relaxed opening angle $\alpha_i$, the energy of a fluctuation with

opening angle $\alpha_i + \Delta\alpha_i$ is well approximated by a quadratic, so that

$$\Delta E(\Delta\alpha_i) \approx \frac{1}{2} A_i (\Delta\alpha_i)^2. \tag{2.11}$$

We can therefore determine the elastic energy barrier for a vertex to have an angle favorable for either insertion ($\alpha_i + \Delta\alpha_i \approx \pi/3$) or joining ($\alpha_i + \Delta\alpha_i \approx 0$), and thus the transition rates, $k_{\mathrm{I}}(\alpha_i)$ and $k_{\mathrm{J}}(\alpha_i)$, respectively. It is now clear that the rates defined above in (2.9)–(2.10) are merely Arrhenius factors, with

$$\sigma^2 = \frac{k_{\mathrm{B}} T}{A_i}. \tag{2.12}$$

Note that $T$ here is temperature, and should not be confused with the triangulation number defined earlier.

During any growth step, new bonds are formed. We may consider an extra energy term, $\mathscr{H}_{\mathrm{bind}} = -N_{\mathrm{b}} E_{\mathrm{b}}$, contributing a binding energy $-E_{\mathrm{b}}$ for each of the $N_{\mathrm{b}}$ bound edges in the capsid. Such an energy is independent of the positional configuration. For our irreversible model to satisfy detailed balance, we need $E_{\mathrm{b}} \gg \Delta E(\Delta\alpha_i)$ so that the energy barrier for the reverse transition is large compared to that for the forward transition.

The parameter $A_i$ and therefore $\sigma$ depends not only on the elastic constants, but also on the local environment of the vertex in the capsid. We can determine normal values for $A_i$ by varying angles on different capsids with different energy parameters, and twice differentiating the Hamiltonian about the minimum. Because most of the opening angle fluctuations in physical situations are *in-plane*, $A_i$ depends most strongly on the Young's modulus, and generally

$$A_i = \frac{\partial^2 \mathscr{H}}{\partial \alpha_i^2} \approx 0.1 \widetilde{Y} r_0^2. \tag{2.13}$$

(Note that this is an absolute dependence on the energy scale $\widetilde{Y}$, and is the only reference we will make to an absolute energy, since everything else depends only on the

ratio $\widetilde{\kappa}/\widetilde{Y} = \ell_{\mathrm{f}}^2$.)

We can perform a rough estimate of this width $\sigma$. Using the elastic parameters estimated for *fd* in §2.1.2, and assuming $r_0 \approx 4\mathrm{nm}$, we find $A_i \approx 0.1 Y r_0^2 \lesssim 17\mathrm{eV}$. We therefore expect fluctuations of

$$\sigma = \sqrt{\frac{k_{\mathrm{B}} T}{A_i}} \gtrsim 0.038 \approx 2.2^\circ \tag{2.14}$$

at room temperature. We will need $\sigma \gtrsim 10^\circ$ for satisfactory growth—a reasonable possibility considering that we conservatively ignored bending rigidity and charge. Had we performed this estimate using the much larger value of $Y$ from $\phi 29$, we would find fluctuations an order of magnitude smaller, leading to a regime in which growth is not feasible. But the $\phi 29$ measurements were taken from the head of a mature bacteriophage that is observed to be much more faceted (small $\ell_{\mathrm{f}}$, large $Y/\kappa$) than the immature form in which assembly occurs. Such small fluctuations are probably important for stability and infectivity, but also quite detrimental to growth[3]. As such, we expect the immature capsid to have much larger fluctuations, although no mechanical studies have been done to allow this determination.

Sometimes a deterministic growth rule is preferred to the stochastic rule presented above. One possibility is a rule that accepts only the move with the largest rate at any given point.

## 2.2 Failure Modes

The restriction that all vertices have either five or six triangles can lead to problems in irreversible growth. It is entirely possible for a growing capsid in our model to perform a wrong growth step resulting in a state that can never be completed—that

[3]For a discussion of the geometric considerations in phage head assembly, see Moody [28]

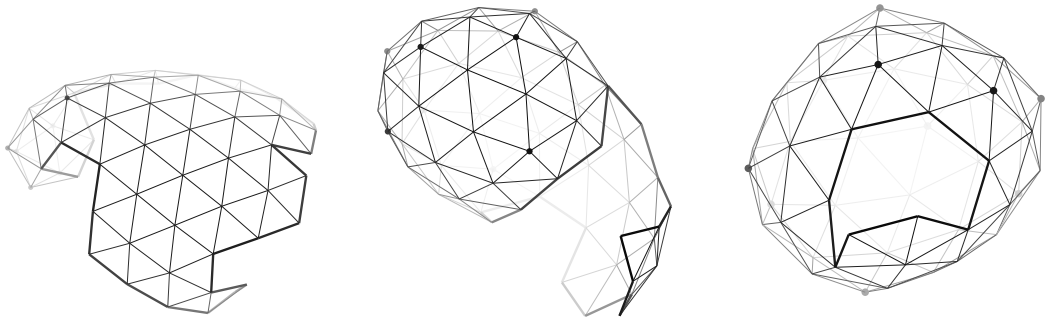


Figure 2.4: A gallery of several failed capsid assemblies.

is, no complete capsid satisfying the pentamer/hexamer-only requirement includes the particular incomplete capsid in any of its possible growth histories. A sample of some failures is shown in Figure 2.4. This section surveys two common failure modes. A common theme is that the failure can be identified non-locally, long before a step is reached at which the growth rules break down; a more rigorous treatment is given in §2.5.

We cannot avoid considering failures, since we must exclude them when reporting statistical distributions of the resulting capsid ensemble (see §2.3). More importantly, we have taken for granted that actual physical assembly has a high success rate (say, 10% to 99%). Indeed, most of our labor on the project reported in this chapter consisted of locating the region of parameter space in which assembly had a high success rate. Classifying the failure modes is a prerequisite to understanding what conditions reduce or eliminate then.

Failure modes are also experimentally pertinent. Whatever the "ideal" capsid is for a given virus species, there is likely to be more than one possible assembly model that produces it. But since different models will tend to fail in different ways, they are better distinguished experimentally by study of defective rather than of ideal capsids. If there are virus species that grow their capsids near the limit of complete irre-

versibility, the resulting ensemble is bound to contain mistakes. Indeed, HIV cones have been observed that are surrounded with what is believed to be a second complete sheet of capsid protein [26].

### 2.2.1 Unfillable quadrilateral hole

First we look at a failure that occurs only at the end of a growth process. Figure 2.5(a) shows a common configuration with a single quadrilateral hole. Parallel transporting a vector around the border gives no rotation and therefore there is no net disclination inside (the net "disclination charge" is zero—recall the discussion in §2.1.1) The only conceivable filling is with two triangles, but either possibility introduces a 7-coordinated vertex[4].

A less trivial example of this situation is shown in Figure 2.5(b). Here we can parallel transport a vector around the border to see that a single disclination must reside within the border; however, there is no way to fill in the remaining triangles to satisfy this. See §2.5 for a more rigorous discussion of this phenomenon. If we continue growth, the hole will eventually shrink to something similar to Figure 2.5(a). Some believe that such a hole is not detrimental to capsids, and in fact capsids are suspected to be permeable to water and ions. On the other hand, HIV is known to have a particle-to-infectivity ratio on the order of 100 [29], and such holes, if they are very common and detrimental to infectivity, may explain why 99% of virions are not infectious.

[4] A 7-fold vertex has a negative disclination charge; there must be a 5-fold vertex next to it, with its positive disclination charge so the interior of the loop is neutral. A pair or "dipole" of positive and negative disclinations constitutes a *dislocation*. Its presence could have been concluded by the non-zero Burgers vector associated with that same loop; that is, the sum of the lattice displacements of each step, referenced to the ideal triangular lattice that can exist far away from this hole

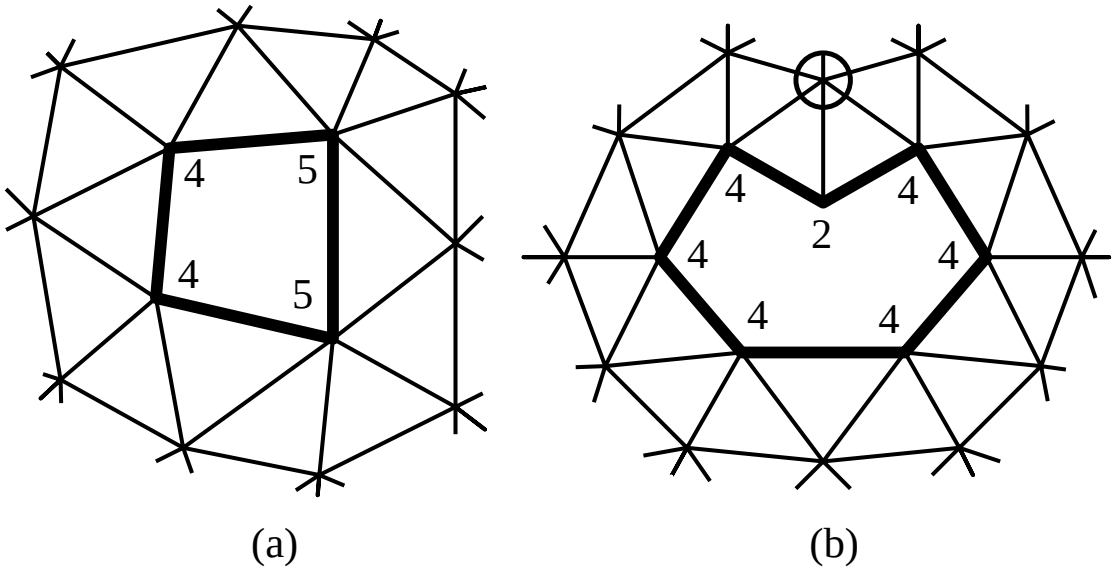



Figure 2.5: Incompletable holes at the end of growth. Neither hole can be filled without introducing a heptamer. The hole in (b) would have been avoided had the circled vertex been made a pentamer

This type of failure was common in all the growth rules we considered, although it is more prevalent in certain situations. In particular, if the growth rate parameters defined in Eqs. (2.9)–(2.10) are large, $\sigma \gtrsim 20^\circ$ or $\Gamma_{\mathrm{I,J}} \gtrsim 200$, then creation of pentamers becomes very random and is no longer based on the configuration. In normal growth, particularly at small $\ell_{\mathrm{f}}$ (angular regime), local strains cause angles along the border to suggest whether a pentamer or hexamer should be created, but large $\sigma$ decreases the sensitivity to this.

### 2.2.2 Crevice formation

Next we look at a failure that can occur at any point during the growth, called a *crevice*. We see in Figure 2.6(a) a portion of a border with the four labeled vertices in a characteristically incompletable configuration. This becomes clear when the border is flattened onto a reference lattice, as seen in Figure 2.6(b). We now see that in the ab-

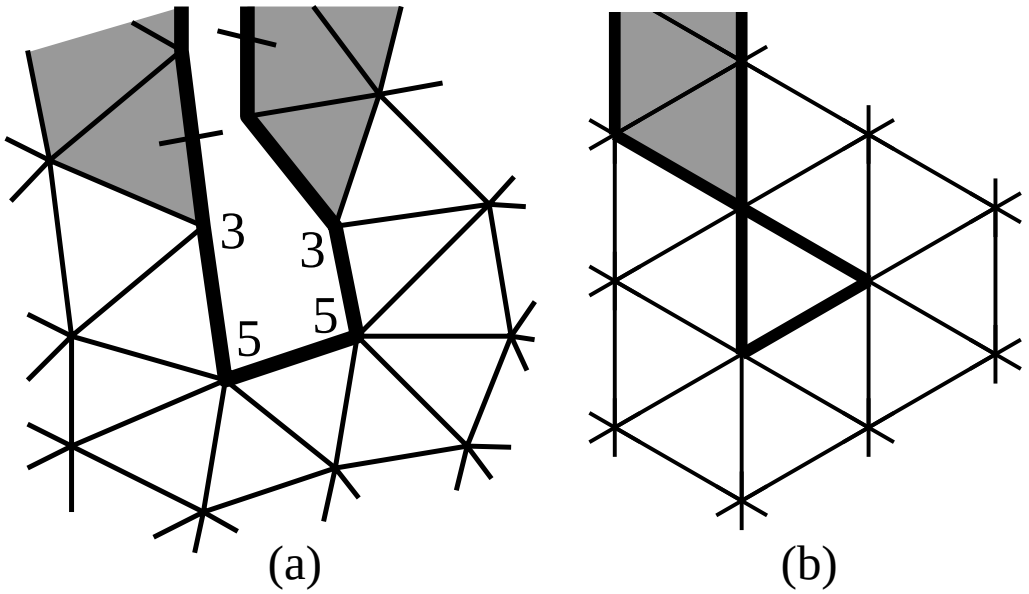



Figure 2.6: (a) An incompletable configuration. The sequence of vertices with 3, 5, 5, and 3 triangles on a border can never lead to a valid complete capsid. Joining the marked edges would produce an unfillable quadrilateral hole, similar to Figure 2.5(a). (b) The same border flattened onto a triangular reference lattice. The shaded triangles from (a) now overlap the corresponding triangles from the opposite side.

sence of pentamers in the neighborhood of this section of border, several triangles lie on top of others. The introduction of a pentamer can only make matters worse. By effectively cutting out a 60° section of the plane, it becomes even more crowded. The only way to alleviate this self-intersection is by introducing a negative disclination (heptamer).

Crevice failures can occur in different regimes, but arise in particular during *fingered* growth. If accretions are much more common than both joinings and insertions (such is the case when either $\sigma \lesssim 5°$ or $\Gamma_{I,J} \lesssim 1$), then we expect many long fingers only one or two triangles wide. Crevices occur easily between these fingers. Even in the absence of fingers, sometimes creating a pentamer will distort a neighboring vertex enough that the angle is too large for insertion or joining. This too often results in a crevice.

Say a single crevice failure occurs during growth. Further growth outward from the failure should be prevented by steric hindrance. But growth elsewhere along the border will continue and eventually fill in the crevice from its far end. Then the capsid will almost complete, leaving a small hole of the same type discussed in §2.2.1 above. The two marked edges in Figure 2.6(a), for instance, might eventually join, leaving a quadrilateral hole.

If two or more crevice failures occur, however, our model's resulting capsid will have a network of cracks connecting these failures. Real capsids might repair this problem by binding edges not sharing a vertex (which is forbidden in our model); in that case, the final result might instead have several small quadrilaterals of the type seen in Figure 2.5(a).

Although it is not as obvious, the smaller holes presented in §2.2.1 also have borders whose flattened images intersect themselves when cut at certain places. We can generalize this by stating that a border is incompletable if there is any choice of cut that leads to any triangles along the flattened border intersecting one another. The converse is true in most cases as well.

### 2.2.3 Failure rates

If 5-fold vertices were simply incorporated at random moments during the growth, virtually every capsid would fail in one of the two modes described in this section. Since the topological constraints to be satisfied are nonlocal, and the growth rates depend on local properties, it seems mysterious at first how the growth can be as successful as it is. The key is that, in an elastic medium, the strain due to a defect (such as a disclination) is also nonlocal; at least, it decays as a power law with the distance from the defect. In this fashion, the necessary information about the location

of a faraway disclination is passed to the growth border.

Since growth is stochastic, there is a possibility of errors despite this passage of information. All capsids are in danger of making an error after the eleventh disclination is in place, and many are in danger even earlier.

We can model the failure probability with a very simple assumption: each time a triangle is added, there is a fixed probability $p_c \ll 1$ of starting a crevice. This is not intended quite literally: $p_c$ must be understood as the fraction of edges along the border that can possibly start a crevice, multiplied by the probability on each such edge that this "wrong" step will be taken when a triangle is added there. (The crucial step might be a "joining" but this contribution gets folded in with the other one, since the border settles into a dynamic near-steady state, so that the ratio of step types will be uniform on average.)

The survival probability of a defect-free capsid is thus

$$\frac{\mathrm{d}P_{\mathrm{sur}}}{\mathrm{d}N} = -p_c \tag{2.15}$$

where step $N$ plays the role of time, so that

$$P_{\mathrm{sur}}(N) = P_0 \mathrm{e}^{-p_c N}. \tag{2.16}$$

Growth will terminate after all twelve disclinations have been incorporated, i.e. on average when $N = \overline{N}(\theta_0, \ell_f)$ (the mean size of capsids formed as a function of the parameters). Furthermore, we hypothesize that $p_c \approx p_c(\ell_f)$, i.e. crevice formation depends strongly on the ratio of elastic constants almost not at all on the preferred angle $\theta_0$. If so, the probability of success is

$$P_{\mathrm{succ}} = P_{\mathrm{sur}}(\overline{N}) = \mathrm{e}^{-p_c \overline{N}}. \tag{2.17}$$

Indeed, we will see the dependence of $p_c$ and $P_0$ on $\ell_f$ in §2.3.2

## 2.3 Results

Here we discuss several measurements that can be used to quantitatively characterize various properties of capsids (individually, or as an ensemble) specified by a triangulation of vertices, such as the results of our growth model. Our results fall into three general categories: size, success, and shape. First we look at the size of the resulting capsids and show the dependence on the elastic parameters. Next we look at the probability of successful growth, in terms of both the size of the capsid and of the growth rate parameters. Last we comment on measures of capsid shape that, along with capsid size, are measurements that can be used with data from cryo-EM experiments.

Figure 2.1 shows an example of a typical capsid shell resulting from our growth simulation. This capsid emphasizes that our configurations are inherently random and irregular. The degree of "lumpiness" in the external shape depends strongly on the Foppl-van Kármán length $\ell_{\mathrm{f}}$, as is elaborated in §2.3.3, below.

Each capsid is grown until either a successful completion, or an identifiable failure, such as a self-intersection in the flattened border. Relaxations are minimized until the gradient-squared is less than $10^{-6}$, in units with $K_{\mathrm{stretch}} \gtrsim K_{\mathrm{bend}} = 1$. The entire growth process for a small capsid takes several minutes on a 1.6GHz processor, while a large capsid takes many hours, the majority of the time devoted to minimizing energy. The following plots of size and success rate include data from 134,352 capsids.

### 2.3.1 Size

The simplest thing to observe about a capsid is its size. We can count the number of triangles $N$, or measure the average radius $R$. As expected [18, 2], capsid size depends most heavily on two parameters from our effective Hamiltonian, $\ell_f = \sqrt{\kappa/Y}$ and $\theta_0$, which we rewrite as a length

$$\ell_\theta \equiv \frac{r_0}{2\sqrt{3}} \cot(\theta_0/2). \tag{2.18}$$

This length is the radius of curvature from two equilateral triangles with side length $r_0$ joined at an angle $\theta_0$ and tangent to a common sphere. We now have three length scales, $r_0$, $\ell_\theta$, and $\ell_f$. It is useful to think of these as two dimensionless parameters, $1/\ell_\theta$ and $1/\ell_f$, taking $r_0 = 1$.

In the smooth regime, when $1/\ell_f \lesssim 50/\ell_\theta$, the variation in the dihedral angles at different bonds is small, so the radius of the resulting capsids tends to $\ell_\theta$. For larger $1/\ell_f$ (the angled regime), the Young's modulus increases. Hexamers, which make up most of the capsid, become flatter. Thus, the effective preferred angle $\theta_0^{\text{eff}}$ decreases, resulting in larger capsids.

We simulated many capsids assembling at four values of $1/\ell_f$ and $\theta_0$ between 7° and 36°. For each set of parameter values, we averaged the radius of the completed capsids, $\overline{R}$, and plot the inverse of the radius $1/\overline{R}$ as a function of the spontaneous curvature $1/\ell_\theta = 2\sqrt{3}\tan(\theta_0/2)$. in Figure 2.7 for several different values of $\ell_f$. We see that, for large $\ell_f$, the curves roughly follow the line $1/\overline{R} = 1/\ell_\theta$. As $\ell_f$ decreases, we see a very different behavior, which favors small (mostly $T = 1$) capsids for a much larger range of $\theta_0$, before the size suddenly increases very quickly around $1/\ell_\theta \approx 0.7$.[5] We can see what is behind these curves in Figure 2.8, which shows the average growth

[5] See §3.2.3 for the same phenomenon in equilibrium.

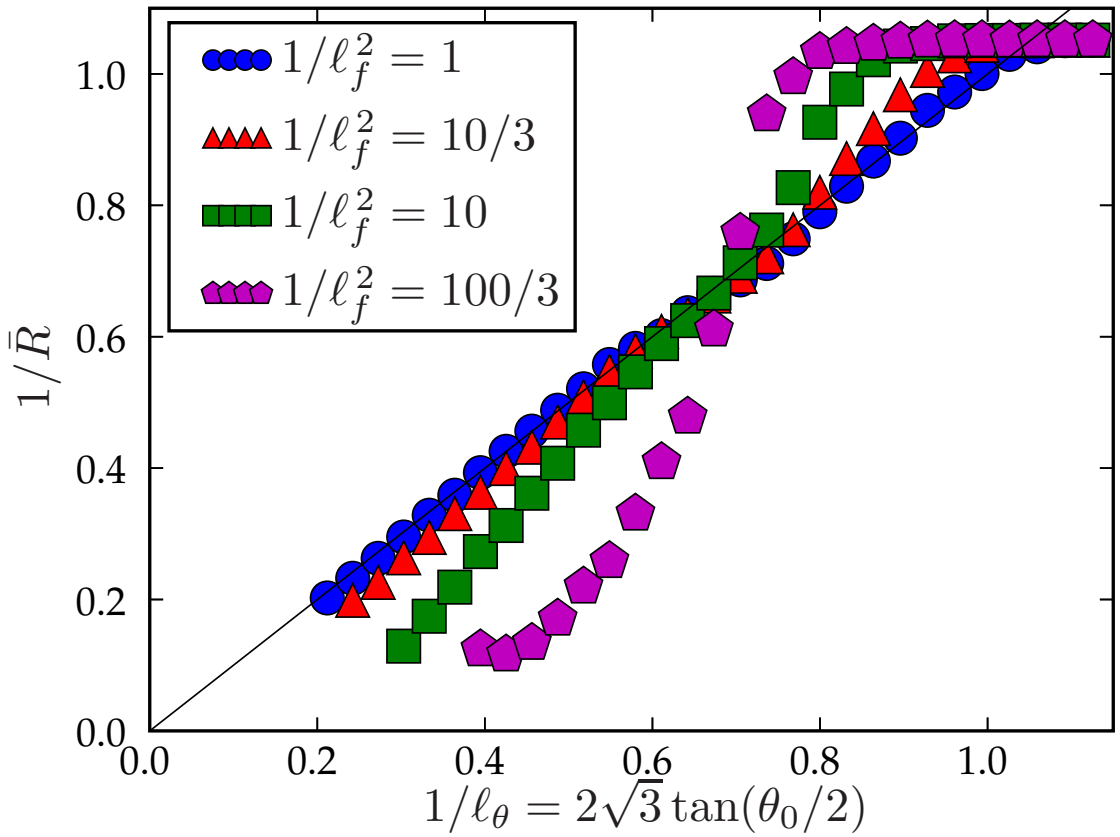



Figure 2.7: Plot of $1/\overline{R}$ versus $1/\ell_\theta = 2\sqrt{3}\tan\theta_0/2$, in units of $r_0$. We see that in the smooth regime of small $1/\ell_\mathrm{f}$, the mean capsid radius $\overline{R}$ very nearly follows $\ell_\theta$. In the angled regime (large $1/\ell_\mathrm{f}$), we find smaller capsids (many $T = 1$) for a much larger range of $1/\ell_\theta$, followed by a sharper increase in size at smaller curvatures. Parameters with fewer than 10 successful capsids were omitted.

history for several individual parameters, represented by the average number of pentamers $\overline{P}$ as a function of the number of growth steps $t$.

### 2.3.2 Success rate

An important consideration for an irreversible growth model is under what circumstances it successfully produces complete capsids. We have already shown that a variety of failure modes exist, resulting in incompletable capsids. We can easily quantify how often these failures actually occur as a function of parameters. We predicted in

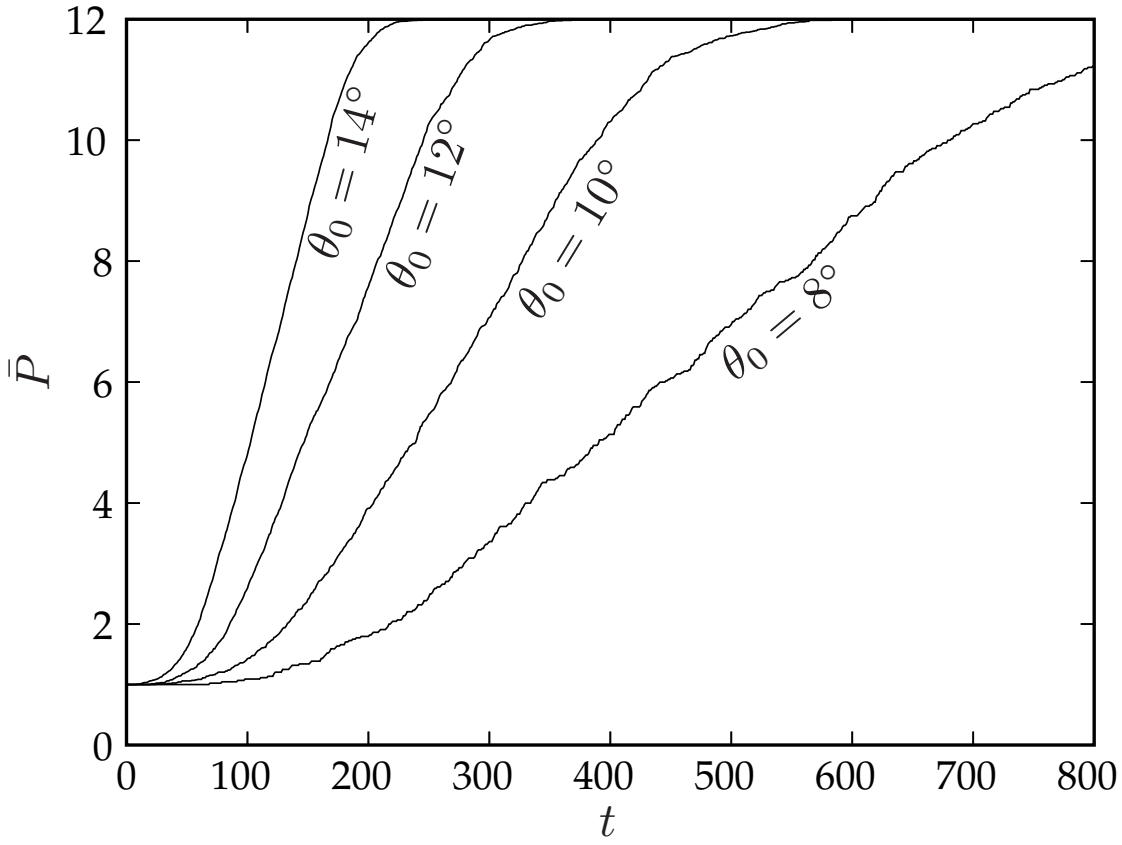



Figure 2.8: Plot of the average number of pentamers $\overline{P}$ versus the step number $t$, which is nearly equivalent to the number of triangles $N$. This gives a picture of the general pathway of growth behind the curves in Figure 2.7. This growth was carried out at $1/\ell_{\mathrm{f}}^2 = 10/3$, and different spontaneous curvatures $\theta_0$ as shown in the figure. We see that growth consists of an initially slow process to add the second pentamer, followed by a rather linear regime in which $\mathrm{d}\overline{P}/\,\mathrm{d}N$ is roughly constant. Note that both the initial rate at $P = 1$ and the following slope depend on $\ell_{\mathrm{f}}$.

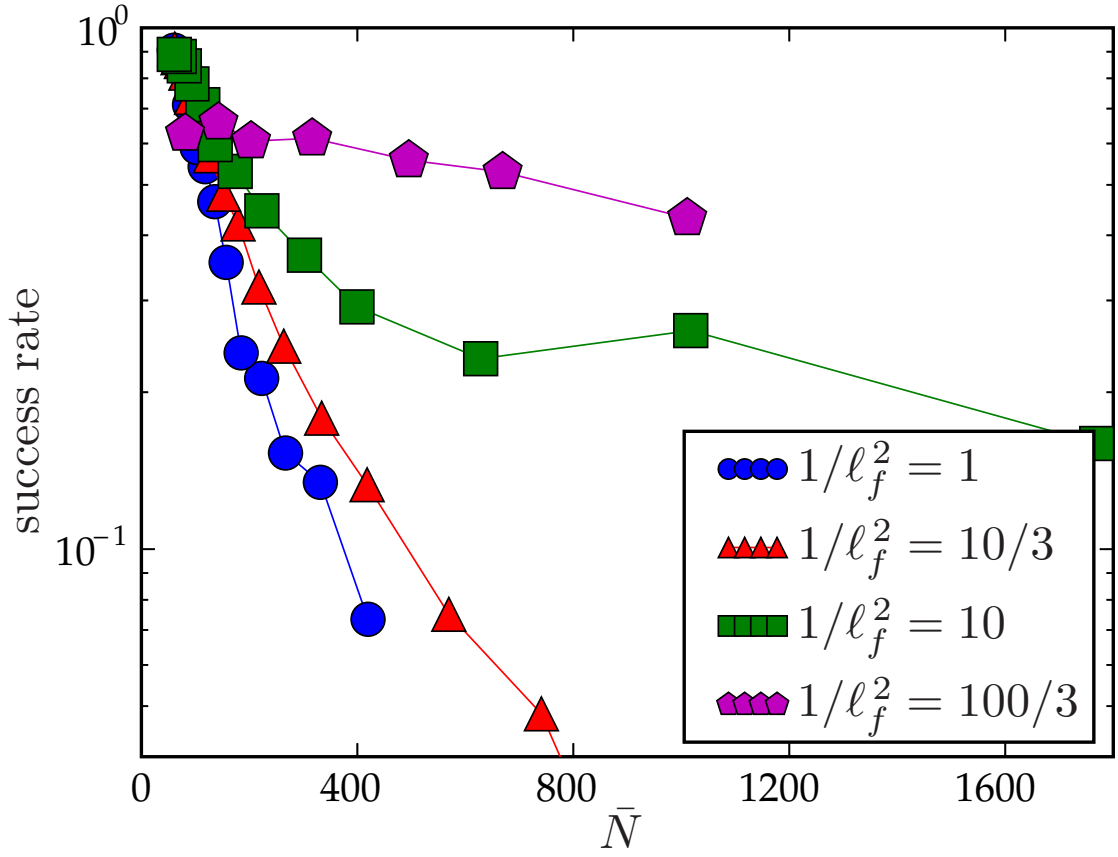



Figure 2.9: Plot of success rates as a function of $\overline{N}$ for the given parameters, from the size measurements. We see a somewhat exponential decay, suggesting that introduction of errors is a Poissonian effect, as discussed in §2.2.3.

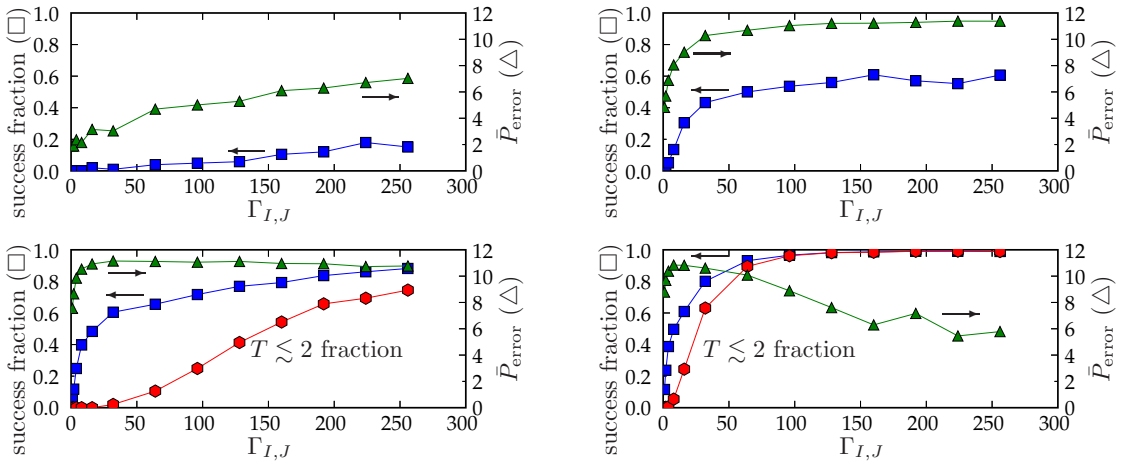



Figure 2.10: Success fractions as a function of rate parameters $\Gamma_{\text{I,J}}$ and $\sigma$ ($\sigma = 8°$, top left; $\sigma = 12°$, top right; $\sigma = 16°$, bottom left; $\sigma = 20°$, bottom right) The square markers show the fraction of successful capsids at each parameter, including small capsids. For $\sigma = 16°, 20°$, the hexagons mark the fraction of total capsids that were small (successful and less than 46 triangles, or pseudo-$T \lesssim 2$). The smaller values of $\sigma$ had no such capsids. Finally, the triangular markers show the average number of pentons $\overline{P}_{\text{error}}$ at the time of error detection. Note that $\overline{P}_{\text{error}} \to 12$ means that all the errors are small holes at the end of growth.

§2.2.3 that the success rate should be exponential with the expected size of the capsid. For each choice of parameters, we average the sizes (measured by number of triangles, in contrast to radius as in Figure 2.7) of the capsids, and thus map the parameters $\theta_0$ and $\ell_{\text{f}}$ to $\overline{N}(\theta_0, \ell_{\text{f}})$. We then plot the percentage of capsids that completed successfully when grown with these parameters in Figure 2.9. While there is systematic deviation from exponential decay, due to the many considerations left out of our analysis, we do still see a mostly exponential trend in the data.

We see in Figure 2.9 that for a given size, growth is generally more successful for more faceted capsids. For large capsids (best-fit radius $\overline{r} \gtrsim 10r_0$), the failures in the smooth regime (i.e. $1/\ell_{\text{f}}^2 \leq 10/3$) all occurred in the early stages of growth, in which only a few pentamers had been added. This suggests that large bending stiffnesses lead to more common crevice failures. On the other hand, the faceted capsids (i.e.

$1/\ell_f \geq 10$) failed mostly in the late stages, in which only several pentamers were missing, suggesting that faceted capsids are somehow resistant to crevice failures and instead fail with small holes.

In §2.2, we mentioned the impact of the rate parameters $\Gamma_{I,J}$ and $\sigma$ on successful completion. We measured the failure rate as a function of these parameters, using reasonable values of $\theta_0 = 16°$ and $\ell_f^2 = 0.1$ In Figure 2.10 we plot the fraction of failed capsids due to either small holes at the end of growth, or crevice failures in the middle of growth. We see that small values of $\Gamma_{I,J}$ and $\sigma$ indeed produce errors. Larger values of $\sigma$ and $\Gamma_{I,J}$ produced successful capsids, but almost all were $T = 1$. This particular result is very sensitive to our particular growth rules, and a choice that prevented insertion until there were five triangles around a vertex would drastically change the result.

### 2.3.3 Shape

Beyond size and success, most other measurements fall under the category of shape measurements. In particular, we might measure either the degree of symmetry or the facetedness of a capsid.

Spherical harmonics may be useful for evaluating icosahedral symmetry, as spherical harmonic coefficients of icosahedrally symmetric functions vanish for all but $\ell = 0, 6, 10, 12, \ldots$ (i.e. $6a + 10b$ for $a, b \in \mathbb{Z}_{\geq 0}$).

Kingston et al. [30] uses the asphericity, defined as the ratio of inradius to circumradius to measure the faceted shape of RSV capsids. Lidmar et al. [4] also defined an asphericity, $\langle R^2 \rangle / \langle R \rangle^2$. While these are good measurements for symmetric capsids, they are not useful for the irregular capsids we grow, because they cannot distinguish between, for instance, a smooth egg-shaped capsid and a faceted spherical capsid.

We instead use a measure based on the Gaussian curvature, described below.

#### 2.3.3.1 Curvature

In light of recent advances in tomography, a very relevant measure is Gaussian curvature $K$. In our discrete triangular model, we can measure the integrated Gaussian curvature $I = \int K \, \mathrm{d}a$ over the neighborhood nearest to a single vertex by measuring the area (equivalently, angle surplus) of the spherical polygon traced out by the incident triangles' unit normals. We can easily extend this to the integrated curvature over all the vertices within any loop around the capsid. The integrated curvature over the entire capsid is always $4\pi$, a topological invariant related to the Euler characteristic. The question then arises how this curvature is distributed over the capsid. For highly faceted capsids, each pentamer has $I \sim \pi/3$, while the rest of the capsid has $I \to 0$. On the other hand, the curvature is distributed uniformly over smooth capsids. This motivates the definition of an inverse participation ratio (IPR),

$$P = \frac{(\sum_j I_j)^2}{\sum_j I_j^2} = \frac{(4\pi)^2}{\sum_j I_j^2}, \tag{2.19}$$

where $I_j$ is the integrated curvature about vertex $j$. This essentially measures the number of lattice sites the curvature is localized to. The IPR is plotted for a single capsid relaxed to different elastic parameters in Figure 2.11. We see that $P = 12$ at $\ell_\mathrm{f} \to 0$ while $P \to N_\mathrm{vert}$ at $\ell_\mathrm{f} \to \infty$.

This same integrated curvature can be measured on triangulated tomographical data from capsids. The integrated curvature within large loops should be relatively stable even if the Gaussian curvature varies quickly. For an arbitrary loop around a capsid, we will get a contribution of $\pi/3$ from each enclosed pentamer. The loop may then be pulled tighter to pinpoint the location of each pentamer. We simulated this process by growing a large number of random capsids and integrating the curvature

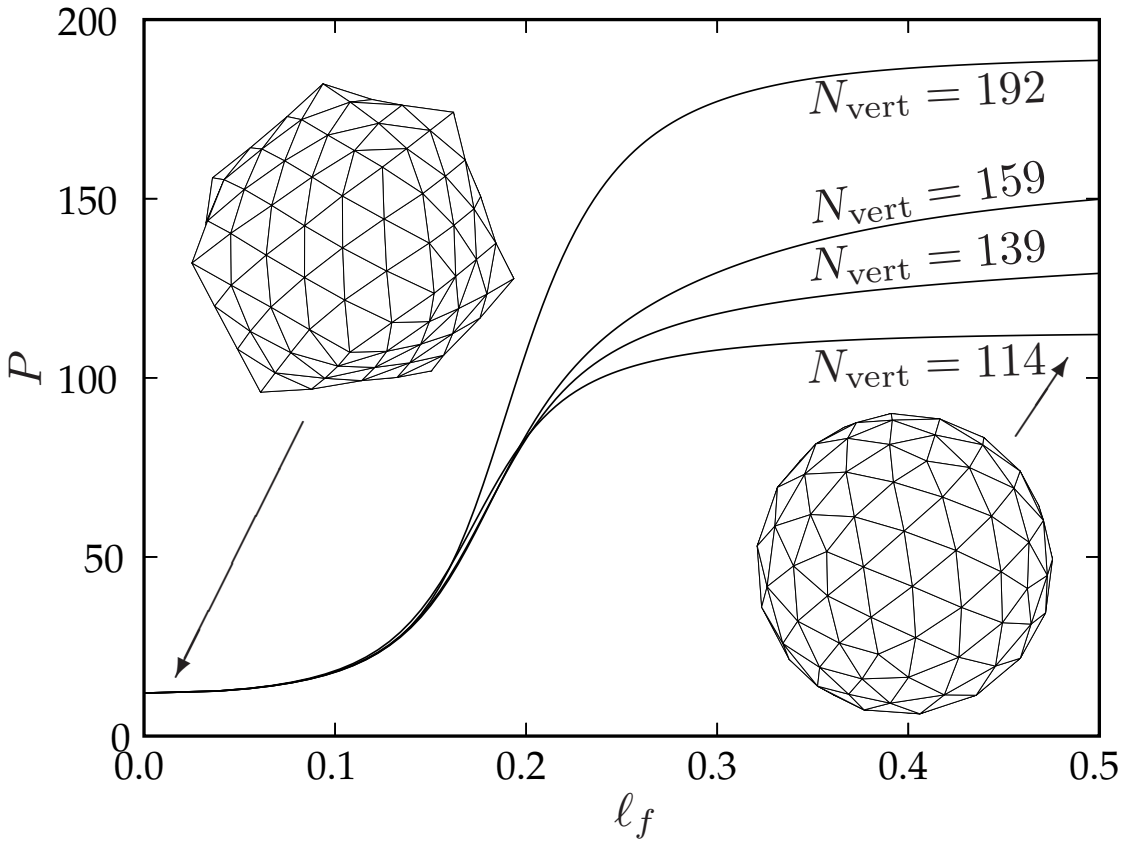



Figure 2.11: Inverse participation ratio for four capsids with different numbers of vertices $N_{\mathrm{vert}}$. We can see that at $\ell_{\mathrm{f}} = 0$ (the angular limit), $P = 12$, and as $\ell_{\mathrm{f}}$ increases (the smooth limit), $P \to N_{\mathrm{vert}}$. The exact shape of the curve depends on the placement of pentamers, but in general we see an inflection point around $\ell_{\mathrm{f}} = 0.18$, which corresponds to $\gamma = R^2/\ell_{\mathrm{f}}^2$ roughly between 200 and 400.

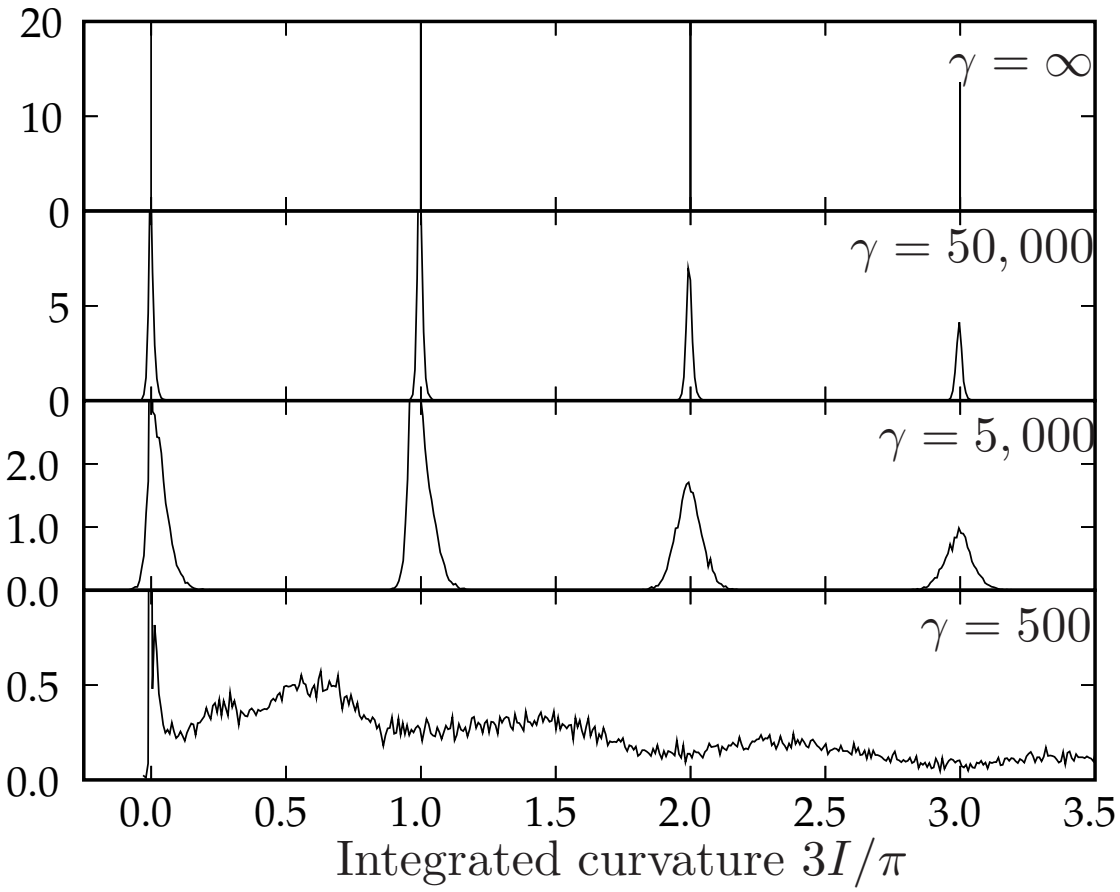



Figure 2.12: Distribution of integral curvatures within random loops around random capsids relaxed to four different elastic parameters, characterized by $\gamma = (R/\ell_f)^2$. The distribution is sharply peaked at the integers for the angled regime at small $\ell_f$ and diffuse for the smooth regime at large $\ell_f$.

within many random loops on each. Each capsid was relaxed to several different values of $\ell_f$. The resulting distribution of curvatures is displayed in Figure 2.12. At large $1/\ell_f \approx 20$ we see very sharp peaks. These peaks diffuse into a mostly uniform background by $1/\ell_f \approx 2$.

#### 2.3.3.2 Average dihedral angle

We can measure the average dihedral angle of either a growing or a complete capsid. Figure 2.13 shows a graph of the average dihedral angle for a very large pentagonal

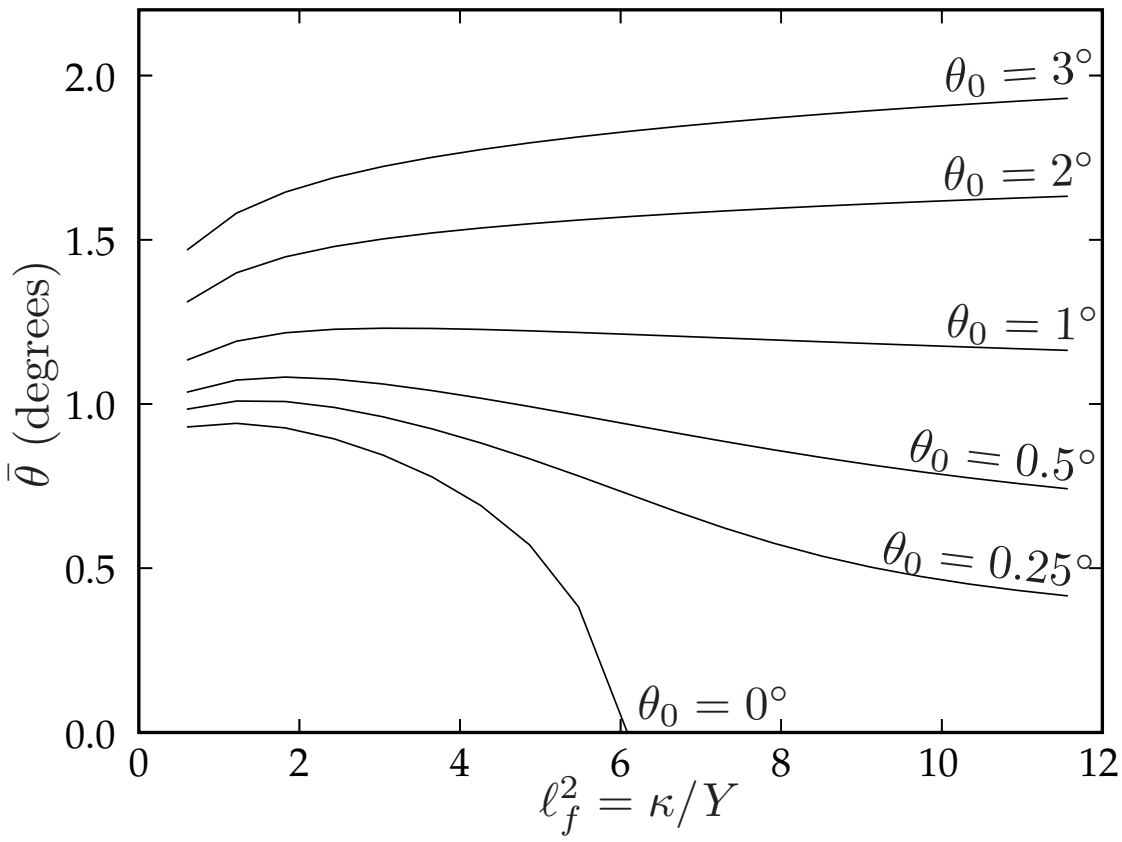



Figure 2.13: Average dihedral angle of a large pentagonal sheet with a single disclination in the center. Plotted versus $\ell_{\mathrm{f}}$ at different $\theta_0$. The bottom $\theta_0 = 0$ curve shows a first-order buckling phase transition. Each subsequent curve increments $\theta_0$ by 0.25°.

sheet with a single disclination in center as a function of $\ell_{\mathrm{f}}$, at different $\theta_0$. We see a first order phase transition at $\theta_0 = 0$.

## 2.4 Discussion

In this section, we recapitulate the highlights of our model and the simulation results, and outline extensions that could improve their realism.

### 2.4.1 Summary of results

Our irreversible growth model, based on trimer units with the simplest possible Hamiltonian and growth rates, did succeed at producing closed capsids, but only when the parameters are tuned to the proper range: $1/\ell_f \lesssim 10$, $\theta_0 \gtrsim 8°$, $\Gamma_{I,J} \gtrsim 50$, and most importantly, $\sigma \approx 12°$.

Our model (§2.1) made a sharp division between configuration variables that were continuous (position) and discrete (bonding topology); correspondingly the model parameters were divided between a Hamiltonian (harmonic form) and rate constants for a set of first-order processes; for simplicity, monomers in solution were not treated explicitly. The model's most distinctive feature is its use of trimer units (triangles), which turns out to have several inherent disadvantages. First, our "insertion" step (§2.1.3) seems as though it should be been redundant; unfortunately, omitting this (relying on the "joining" step in its stead) produces abundant failures—the fingered growth and crevices elaborated in §2.2.2. In other words, good growth depended on joinings being rare compared to insertions, which followed from our growth rules (§2.1.3), since opening angles $\alpha$ near 0° are much less common than those near 60°. This may mean that if a capsid assembles from trimers in solution, the only way to have normal growth is that there must be cooperative binding as in our insertion step.

Our results may be divided into two categories: the growth process (including the success rate) and the shape of the resulting model capsids. In the first category, we found mathematical descriptions (§2.5) that clarified the constraints on the positions of the fivefold-coordinated vertices, which fully characterize the bond network. Additionally, we uncovered a simple relation between the chance of failure and capsid size, (2.17). In the second category, we showed the relationship between the capsid's

final size and the two length parameters, $\ell_f$ and $\ell_\theta$. In contrast to Nguyen et al. [23], we note that the *average* capsid size is indeed well-determined by the spontaneous curvature parameters for large capsids, even in the absence of any scaffolding considerations. We also extended the concepts of Lidmar et al. [4] to irregular capsids. In particular the ratio of bending and stretching stiffnesses—which we suggest is best parametrized by a length, (2.5), rather than a dimensionless ratio—controls whether the resulting shape is smooth or angular, as we have characterized by an inverse participation ratio, (2.19).

One application of our results from different $\ell_f$ relates to phage maturation. Many phages include a maturation step in which the assembled prohead greatly increases its stretching stiffness relative to its bending stiffness, making it more faceted. One might ask why the assembly process occurs in the smooth regime, especially since our results in §2.3.2 show that the probability of success is smaller in this regime. We propose that the advantage of growth in the smooth regime is size selection. In Figure 2.7 we see that faceted capsids have a sharp transition in size around the region most phages fall into ($T = 3$ to $T = 7$). As might be expected, for a given set of parameters, the spread of capsid sizes is also much broader near this transition. Thus, in order to well-control the size of assembled capsids, a virus might prefer to grow in the smooth regime, counting on other factors such as scaffolding to increase its chances of successful assembly.

### 2.4.2 Future directions: more realistic random growth

#### 2.4.2.1 Models with non-trimer units?

The retroviral CA proteins we claim to model have well-documented dimerization [31] and hexamerization [32] interactions, but no trimer interactions have been

observed in retroviral capsids. A model based instead on pentamers and hexamers could be implemented simply by changing the growth steps to add several triangles at a time, so as to fully enclose a single vertex each step into either a hexamer or a pentamer. We gain some benefit, however, from actually changing our representation to a honeycomb lattice—the dual to our current triangular lattice. Vertices of the dual lattice are all three-coordinated, so each vertex along the border has either one or two capsomers attached to it—much simpler than the five different possible coordination states for border vertices in the trimer model. In this model, growth rules could explicitly depend on the total coordination of a vertex. Such coordination-based rules greatly assisted successful growth in our trimer model, but were not as physically justifiable as they are in the dual model.

These considerations suggest that behavior arising from this choice is not universal. We expect models based on dimers, trimers, or pentamers and hexamers to fall into different universality classes.

Another direction leading to a more realistic model is to improve the accuracy of our interactions. Microscopic electrostatic simulations, such as with the CHARMM software, could provide a more realistic Hamiltonian for specific viruses, which could be included in future models.

#### 2.4.2.2 Lattice fluctuations

A deeper understanding of the relationship between topological configurations is critical. So far we have only thoroughly considered irreversible growth transitions. Other transitions relate to the motion of disclinations on the lattice (always in pairs), both for the purpose of enumerating the near-symmetric states, and for an understanding of the rearrangement dynamics by which real capsids may anneal their

bond configurations into the free energy minima predicted by many equilibrium models.

### 2.4.3 Future directions: realistic shapes

All well-studied real capsids exhibit greater regularity than our current model can regularly generate. How can the Hamiltonian (or the growth dynamics) be modified so as to generate an icosahedral, or (for HIV) conical capsid?

#### 2.4.3.1 Icosahedral symmetry

The main challenge for theory is to explain the assembly of icosahedrally symmetric capsids, if one is not close to equilibrium. Hamiltonians such as ours do indeed give effective repulsion between the disclinations, and the free energy minimum is known to have icosahedral symmetry in similar models [2, 4]. However, this is simply insufficient to produce large symmetric capsids in a model where the accretion rate depends on local geometry, since the growing border does not contain enough information in just the opening angles (§2.1.3). Even deterministic variants of the growth model never yielded icosahedral capsids larger than $T = 4$.

We speculate that if the bending potential $\mathscr{H}_{\text{bend}}$ was not simply harmonic around $\theta_0$, but instead had minima at two different angles $\theta_1$ and $\theta_2$, this might robustly favor a regular pattern of edges with $\theta_1$ and $\theta_2$, thus permitting determination of larger icosahedral capsids. A double-well potential would presumably represent some sort of conformational switch, perhaps an internal bending between two domains of the capsid protein. Thus, this proposal has some features in common with the matching-rule models that we dismissed as implausible (§1.4), but anharmonic potentials seem much more natural than variations in the edge-binding (which, in

our model, corresponds to the term $\mathscr{H}_{\text{bind}}$ mentioned briefly in §2.1.3).

One other change that could result in more symmetric capsids, as well as more successful growth in general, is to relax our irreversibility constraint. Allowing the growing edge to "melt back" would allow a growing capsid to better explore the possible configurations, in particular curing crevice and fingering defects.

#### 2.4.3.2 Retroviruses

We asserted that the randomness of our model's growth behavior makes it appropriate for modeling the irregularity and pleiomorphism observed in the capsids of retroviruses such as HIV. However, mature HIV capsids do have a typical gross shape, which is mostly conical (although sometimes tubular) in vivo, whereas our current model grows round capsids on average. A cone is characterized by having (say) five disclinations around its smaller end, seven around the large end, and none on the belt between; this means the rates of adding pentons must somehow vary during different stages of the growth. When cones form inside an envelope, the difference could be attributed to depletion of the monomers as they are incorporated into the capsid: that (see §2.1.3) would decrease the rate of insertion but not of joining, leading to a greater chance of penton formation. A difficulty with concentration control is that cone completion leaves in solution 70% [33] of the capsid proteins: in order for this to grossly affect the rates, accretion must microscopically be a rather high-order process. It also leaves unexplained the large density of pentons at the *earliest* stage: a possibility is to add a simple interaction between the capsid and either the nucleic acid or the membrane [26, 34].

## 2.5 Completability

It is possible to grow an incomplete capsid that is not part of any allowed capsid. This appears to be a consequence of our rule that a capsid vertex can only have coordination 5 or 6. (Seven-coordinated vertices, were they allowed, would let the capsid recover from almost every "mistake" discussed in this section.)

As a complement to the more qualitative discussion in §2.2, this section presents the technical criteria we discovered to identify when a partial capsid is or is not completable, non-locally and long before the growth rules carry us to a point where we must make a 7-fold vertex or stop. The completability conditions are defined entirely in terms of the growing border, which can be uniquely described by traversing the vertices (in a specified direction) and listing the number of triangles present at each vertex. Thus a string of numbers from 1 to 5 specifies a border. (6 is allowed, but is trivial.)

### 2.5.1 String representation

We can represent any border by a word $a_1 a_2 \ldots a_n$, where $1 \le a_i \le 6$ is the number of triangles around the $i^{th}$ vertex, counting clockwise from an arbitrary starting point. We define several operations on these string representations. First, consider

$$A(a_1 a_2 \ldots a_n) \equiv 1(a_1+1)a_2 \ldots (a_n+1) \tag{2.20}$$

and

$$J(a_1 a_2 a_3 a_4 \ldots) \equiv (a_1 + a_3)a_4 \ldots, \tag{2.21}$$

representing accretion and joining, respectively. Note that the $a_2$ term disappears upon applying $J$. This vertex is enclosed and is no longer part of the border. We therefore require $a_2 = 5$ or 6. We can further define insertion $I = J \circ A$ as the composition

of joining and accretion. Finally, because we defined these operations to act on the starting and ending points of our string representation, we must define a cycle operation, $C(a_1 a_2 \ldots a_n) \equiv a_2 \ldots a_n a_1$. Because of the unimportance of the starting point in representing a border, cycling leaves borders invariant. Since these operations are sufficient to grow any capsid, we can uniquely describe a capsid by the sequence of operations on the border required to arrive at the border from a single triangle, 111.

Using this representation we can immediately identify some borders as incompletable. Consider the border $X = 555\ldots$ Joining is illegal since it leaves a vertex with 10 triangles. Accretion leads to $A(X) = 6166\ldots$, which clearly cannot be completed since only joining can be done on the 6's, and this leaves seven triangles about at least one vertex. Finally, insertion yields $I(X) = 66\ldots$, which is incompletable for the same reason.

Any border that intersects itself on a flat reference lattice is incompletable (coincident edges are allowed). It is important to take notice of which side of the border is the inside (from which the triangles are being counted) and which is the outside.

We thus define the complement of a border

$$\overline{a_1 \ldots a_n} \equiv (6 - a_n) \ldots (6 - a_1). \tag{2.22}$$

If the original border enclosed $d$ disclinations then its complement encloses $12 - d$ and can be glued together to form a complete capsid. We must note two things. Firstly, the complement of a border may be a border that cannot possibly be grown using our growth operations. Secondly, the complement is only unique insofar as the seam between the two incomplete capsids is occupied only by six-fold vertices. However, many "pseudo-complements" may be constructed that leave disclinations on this seam.

While the border by itself is useful for analyzing completability, it does not

uniquely describe the interior. An individual border may have many different realizations, with disclinations in different positions. In fact, a pair of disclinations can move in opposite directions (relative to a common reference lattice, if one exists) without changing the border.

### 2.5.2 Winding number

We can compute the winding number $W(a_1 a_2 \dots a_n) \equiv \sum_i (a_i - 3)$ of a border, which is the number of 60° turns undergone by a direction that is parallel transported about the border. The net number of disclinations within the border is $W + 6$. If we allowed seven-fold disclinations, they would be subtracted from this number. Since we only allow single positive disclinations, we can conclude that the winding number around any path on a valid capsid must be between $-6$ and $+6$, leaving $6 - W$ disclinations to be placed in the unfilled part (the other side of the border, counting the vertices on the border itself).

### 2.5.3 Six disclinations remaining

We will now show that any border with winding number $W = 0$ that does not intersect itself on a flat reference lattice is completable by applying a finite number of growth operations to the border, resulting in a self-complementary border of the form $3^m 43^n 2$, which can be glued onto a copy of itself to make a complete capsid.

First draw the border on a flat reference lattice. It is now clear that triangles can be added to the border to transform it to the required form. So any capsid with a non-intersecting border and $W \leq 0$ is completable.

This procedure is demonstrated in Figures 2.14 and 2.15.

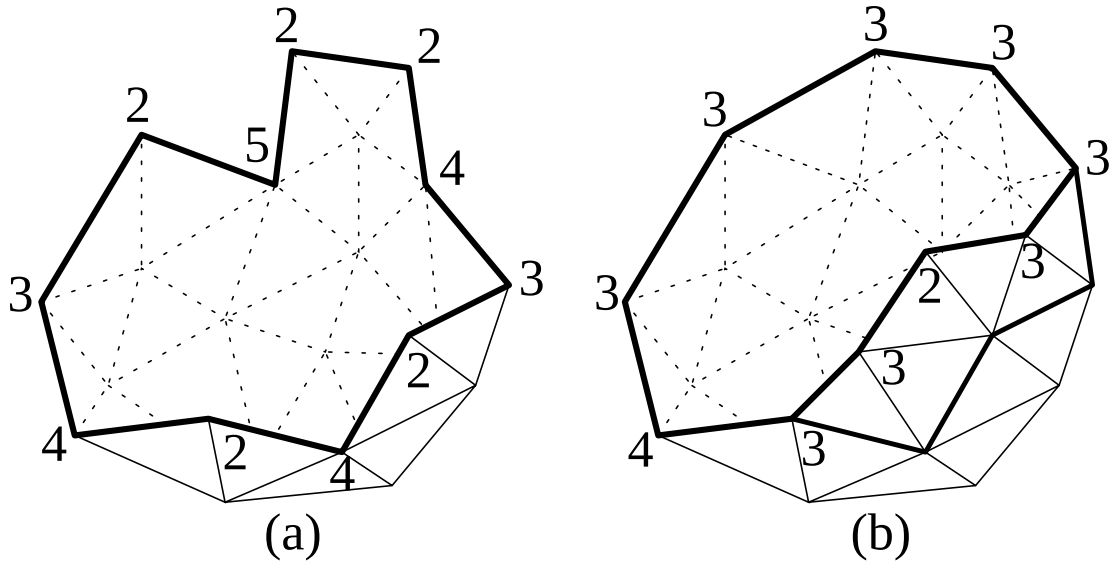



Figure 2.14: Adding triangles to a $W = 0$ border (a) to transform it into the self-complementary form $3^m23^n4$ seen in (b).

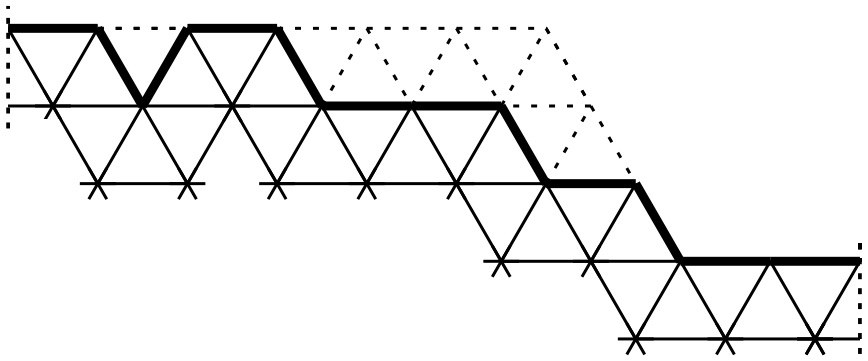


Figure 2.15: An alternate point of view of the same procedure as illustrated in Figure 2.14. We flatten the border onto a flat triangular reference lattice. The dashed lines on the left and right correspond to the place the border has been cut. That $W = 0$ is evident because there is no net rotation after traversing the border. It is also clear that adding the dashed triangles results in the same $3^m23^n4$ border as above.

### 2.5.4 Late stage completability

When $W > 0$, there are more constraints. We can no longer add triangles freely since every row we add is smaller due to the enclosed disclinations. We will begin by considering the case of an incomplete capsid with eleven disclinations enclosed, leaving a deficit of one disclination needing to be placed.

#### 2.5.4.1 One disclination remaining

In this case we can easily look at the reverse picture. If the border is completable then it is a path on a valid complete capsid and we can therefore look for a pseudo-complementary border to fill it. We can represent a triangular lattice with a single disclination as a flat triangular lattice with a 60° section cut out and the edges identified. If we therefore flatten our border onto a flat lattice, we expect the first and last points to be identified by this edge and therefore we can draw an equilateral triangle with the third point at the required location of the disclination. While the edges of the triangle need not be along a lattice direction, the third point is necessarily on the lattice. The border is completable if and only if this disclination is at an unoccupied point (outside of the original border). Note that because the border has a 60° rotation, this point is unique, regardless of the choice of starting and ending point. This process is demonstrated in Figure 2.16.

#### 2.5.4.2 Two disclinations remaining

Two disclinations ($W = 4$) works in a very similar way to the single disclination discussed above, except we have a 120°-30°-30° isosceles triangle instead. This gives a single charge +2 disclination, but since we do not allow two disclinations at the same point, we must move them slightly. Figures 2.17 and 2.18 show the two possible sit-

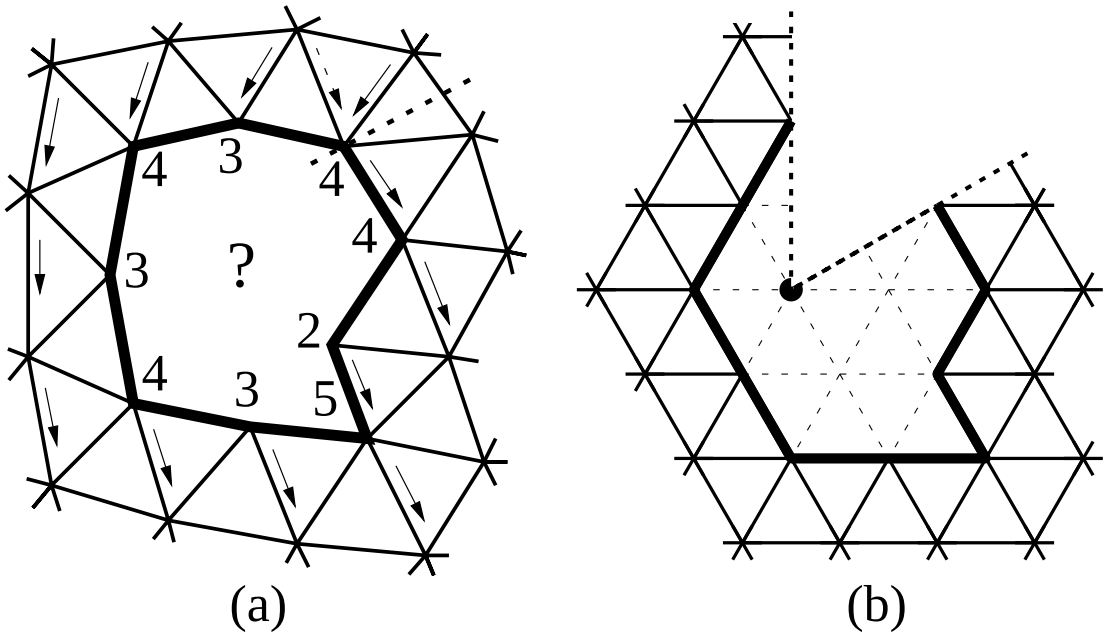



Figure 2.16: A border with $W = 5$. Parallel transporting a direction around the border yields a rotation of 60°. A disclination must therefore be located within this border. The dashed line shows where we plan to cut. (b) We see the same border flattened onto a reference lattice. The 60° rotation is now more clear. The dashed lines are part of an equilateral triangle and therefore show the required location of a disclination. Any choice of cut results in the same location, as long as the of the cut is chosen so that the triangle is equilateral.

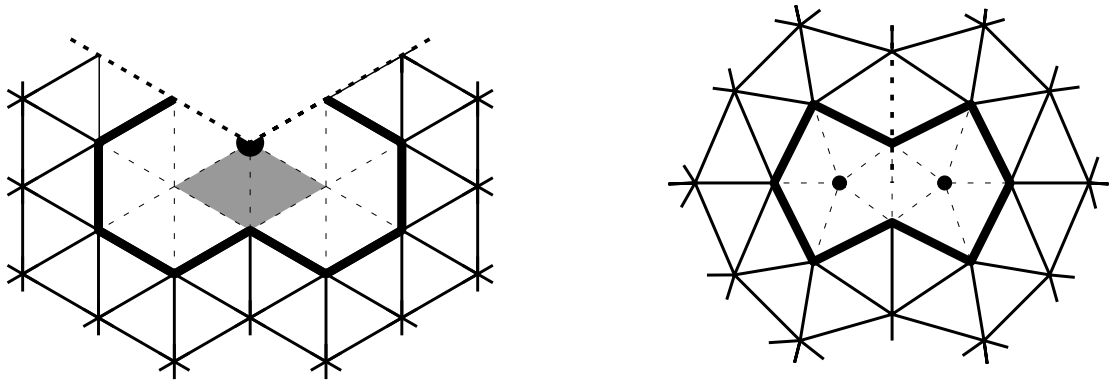


Figure 2.17: Rearrangement of a +2 (120°) disclination located on a vertex into a pair of single disclinations with the same border. The two shaded triangles are removed.

uations and equivalent fillings with only single disclinations and the same border. If the center is on a lattice point, then the disclinations can each move in opposite directions to neighboring points and the same region of the plane will be cut out, up to a triangle at the apex, as shown in Figure 2.17. If the center is in the center of a triangle rather than on a lattice point, we can place the two disclinations on adjacent lattice points around the triangle for the same effect, as shown in Figure 2.18. The disclinations can be further separated in a similar fashion.

This breaks down if the +2 disclination is on a vertex on the border that has 4 or more triangles. In this case there is no way to separate the disclinations without one of them crossing the border.

#### 2.5.4.3 Three disclinations remaining

The case of $W = 3$ follows the same way, except now we find a +3 disclination on the midpoint of a line segment joining the two identified points. This +3 disclination may be on a lattice point or on the edge of a triangle. Both can again be split similarly to the previous case, as seen in Figures 2.19 and 2.20. As seen in the flattened pictures, the +3 disclination is always within the border, provided the flattened border

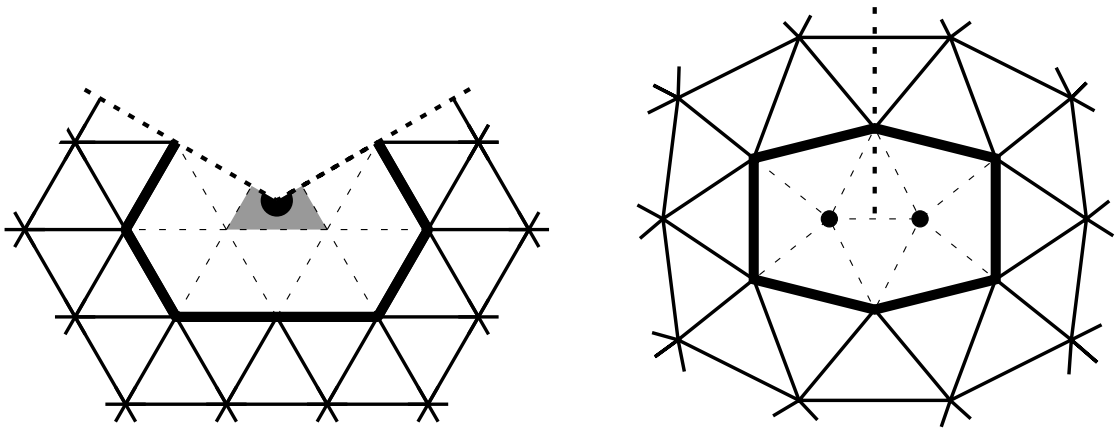


Figure 2.18: Rearrangement of a +2 (120°) disclination located on a triangle into a pair of single disclinations with the same border. The shaded part of the triangle is removed.

does not intersect itself or the "cut line". Thus outside of these cases, the border is only incompletable if the +3 disclination cannot be split properly without any single disclinations crossing a border.

## 2.6 Steric considerations

Our triangular units are two-dimensional objects but they represent three-dimensional structures in space. Thus, we must explicitly ensure that two triangles can never be in positions, such that the proteins they represent would overlap in space. This section collects details concerning the implementation of steric constraints. First (§2.6.1) we write the explicit form of the term in our Hamiltonian that prevents self-intersection; then (§2.6.2) we discuss the way in which steric constraints tend to assist growth and to discourage the wrong steps that lead to failure.

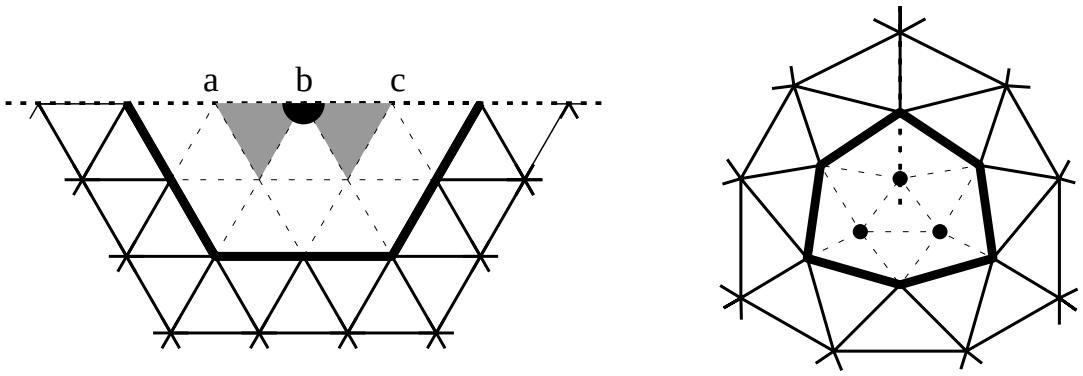



Figure 2.19: Rearrangement of a +3 (180°) disclination located on a vertex into three single disclinations with the same border. The two shaded triangles are removed. Note that vertices *a*, *b*, and *c* all come together to form a single five-fold vertex.

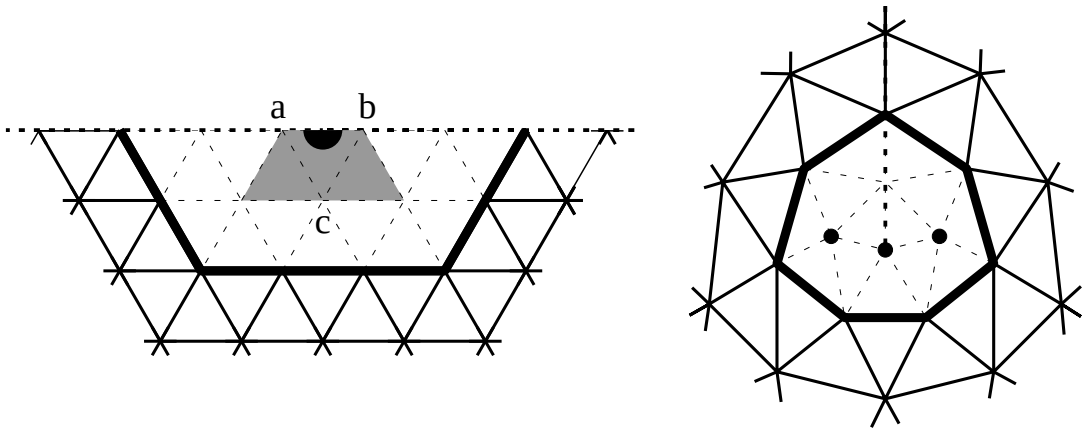



Figure 2.20: Rearrangement of a +3 (180°) disclination located on an edge into three single disclinations with the same border. The three shaded triangles are removed and the vertices *a*, *b*, and *c* collapse to a single five-fold vertex.

### 2.6.1 Steric potential

The final term in (2.1) was a steric repulsion term: since our capsid units are two-dimensional triangles, some such term has to be added by hand, to account for the thickness of our three-dimensional proteins and disfavor unphysical configurations. The details of this term were deferred from §2.1.2 to this section. The steric term should have the simplest possible form, in keeping with the toy-model spirit of our other terms.

In the steric term, the two kinds of degrees of freedom—topological and positional—clash in a sense. Two units that are nearby in space may be many steps apart on the bond network, and thus practically decoupled from each other. (There is little interaction in the elastic energy, and furthermore the ways they constrain the available discrete growth steps are independent.) Hence, $\mathscr{H}_{\text{steric}}$ must consist of topologically long-range, but positionally short-range, interactions.

We chose an implementation based on augmenting each triangle by another vertex over the face (on the interior side), thus forming a tetrahedron. We define a repulsion between the apex vertex of each tetrahedron and every (non-apex) vertex of every other triangle. Thus,

$$\mathscr{H}_{\text{steric}} = \sum_{I,j} V_{\text{steric}}(|\boldsymbol{r}_I^{\Delta} - \boldsymbol{r}_j|), \tag{2.23}$$

where $\sum_I$ is a sum over triangles and $\boldsymbol{r}_I^{\Delta}$ is an equal distance $\ell_{\text{steric}} \lesssim r_0$ inward from the three vertices of the triangle. Furthermore, we require $V_{\text{steric}}(r) = 0$ if $r \geq \ell_{\text{steric}}$, which is the case for all pairs $I, j$ in most capsids. This form allows the edges of unconnected triangles to be incident while maintaining $\mathscr{H}_{\text{steric}} = 0$ so long as the triangles do not actually intersect.

We choose the simplest form that is differentiable at $r = 0$ and $r = \ell_{\text{steric}}$,

$$V_{\text{steric}}(r) = k_{\text{steric}}(\ell_{\text{steric}}^2 - r^2)^2, \quad r < \ell_{\text{steric}}, \tag{2.24}$$

Choosing $\ell_{\text{steric}} \approx 0.65 r_0$ generally provides sufficient stericity while not interfering with the shape of non-self-intersecting capsids.

It is important to stress that this steric term should not affect most capsids. For non-growing capsids, we generally turn it off to increase efficiency, since it always vanishes.

### 2.6.2 Steric growth heuristics

While the steric potential discussed in §2.6.1 is useful to prevent capsids from relaxing to unphysical positions, it does not directly help the growth rules. Because growth rules are based entirely on rates $k_{\text{A}}$, $k_{\text{I}}$, and $k_{\text{J}}$ derived from the local geometry around individual vertices, there is no way to directly determine whether a step will cause a steric hindrance. Because such growth steps are not likely to occur in nature, we implement a heuristic to detect such steps and remove them from the set of allowed growth steps by setting the rate to zero.

Before any accretion or insertion, we perform two tests. First we look at the steric potential $\mathscr{H}_{\text{steric}}$. If the accretion causes $\mathscr{H}_{\text{steric}} \neq 0$ then the accretion fails. Next, if the accretion causes the centroid of one triangle to be within $\ell_{\text{steric}}/\sqrt{10}$ of the vertex of another triangle, then the accretion fails. This is necessary because the first test misses the case where two triangles are directly on top of one another. This case is less important while minimizing, because minimization would need to pass a large energy barrier, while growth steps can jump over it for free.

## BIBLIOGRAPHY


[1] S. D. Hicks and C. L. Henley, "Irreversible growth model for virus capsid assembly", *Phys. Rev. E* **74**, 031912 (2006).

[2] R. F. Bruinsma, W. M. Gelbart, D. Reguera, J. Rudnick, and R. Zandi, "Viral Self-Assembly as a Thermodynamic Process", *Phys. Rev. Lett.* **90**, 248101 (2003).

[3] R. Zandi, D. Reguera, R. F. Bruinsma, W. M. Gelbart, and J. Rudnick, "Origin of icosahedral symmetry in viruses", *Proc. Nat. Acad. Sci. USA* **101**, 15556 (2004).

[4] J. Lidmar, L. Mirny, and D. R. Nelson, "Virus shapes and buckling transitions in spherical shells", *Phys. Rev. E* **68**, 051910 (2003).

[5] D. Endres, M. Miyahara, P. Moisant, and A. Zlotnick, "A reaction landscape identifies the intermediates critical for self-assembly of virus capsids and other polyhedral structures.", *Protein Sci.* **14**, 1518 (2005).

[6] A. Zlotnick, R. Aldrich, J. M. Johnson, P. Ceres, and M. J. Young, "Mechanism of Capsid Assembly for an Icosahedral Plant Virus", *Virology* **277**, 450 (2000).

[7] J. A. Speir, S. Munshi, G. Wang, T. S. Baker, and J. E. Johnson, "Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy.", *Structure* **3**, 63 (1995).

[8] P. Forrer, C. Chang, D. Ott, A. Wlodawer, and A. Pluckthun, "Kinetic Stability and Crystal Structure of the Viral Capsid Protein SHP", *J. Mol. Biol.* **344**, 179 (2004).

[9] Z. Xie and R. W. Hendrix, "Assemblyin Vitroof Bacteriophage HK97 Proheads", *J. Mol. Biol.* **253**, 74 (1995).

[10] U. K. von Schwedler, K. M. Stray, J. E. Garrus, and W. I. Sundquist, "Functional Surfaces of the Human Immunodeficiency Virus Type 1 Capsid Protein", *J. Virol.* **77**, 5439 (2003).

[11] B. Berger, P. W. Shor, L. T. Kellogg, and J. King, "Local rule-based theory of virus shell assembly", *Proc. Nat. Acad. Sci. USA* **91**, 7732 (1994).

[12] D. C. Rapaport, "Self-assembly of polyhedral shells: A molecular dynamics study", *Phys. Rev. E* **70**, 051905 (2004).

[13] A. Zlotnick, "To Build A Virus Capsid - An Equilibrium-Model Of The Self-Assembly Of Polyhedral Protein Complexes", *J. Mol. Biol.* **241**, 59 (1994).

[14] D. Endres and A. Zlotnick, "Model-Based Analysis of Assembly Kinetics for Virus Capsids or Other Spherical Polymers", *Biophys. J.* **83**, 1217 (2002).

[15] A. Zlotnick, "Are weak protein-protein interactions the general rule in capsid assembly?", *Virology* **315**, 269 (2003).

[16] H. S. Seung and D. R. Nelson, "Defects In Flexible Membranes With Crystalline Order", *Phys. Rev. A* **38**, 1005 (1988).

[17] S. Li, C. P. Hill, W. I. Sundquist, and J. T. Finch, "Image reconstructions of helical assemblies of the HIV-1 CA protein", *Nature* **407**, 409 (2000).

[18] D. L. D. Caspar and A. Klug, "Physical Principles In Construction Of Regular Viruses", *Cold Spring Harb. Symp. Quant. Biol.* **27**, 1 (1962).

[19] N. Nandhagopal, A. A. Simpson, M. C. Johnson, A. B. Francisco, G. W. Schatz, M. G. Rossmann, and V. M. Vogt, "Dimeric Rous Sarcoma Virus Capsid Protein Structure Relevant to Immature Gag Assembly", *J. Mol. Biol.* **335**, 275 (2004).

[20] V. S. Reddy, H. A. Giesling, R. T. Morton, A. Kumar, C. B. Post, C. L. Brooks, and J. E. Johnson, "Energetics of Quasiequivalence: Computational Analysis of Protein-Protein Interactions in Icosahedral Viruses", *Biophys. J.* **74**, 546 (1998).

[21] G. A. Vliegenthart and G. Gompper, "Mechanical Deformation of Spherical Viruses with Icosahedral Symmetry", *Biophys. J.* **91**, 834 (2006).

[22] I. L. Ivanovska, P. J. de Pablo, B. Ibarra, G. Sgalari, F. C. Mackintosh, J. L. Carrascosa, C. F. Schmidt, and G. J. Wuite, "Bacteriophage capsids: tough nanoshells with complex elastic properties.", *Proc. Nat. Acad. Sci. USA* **101**, 7600 (2004).

[23] T. T. Nguyen, R. F. Bruinsma, and W. M. Gelbart, "Elasticity theory and shape transitions of viral shells", *Phys. Rev. E* **72**, 051923 (2005).

[24] T. Maeda and S. Fujime, "Dynamic light-scattering study of suspensions of fd virus. Application of a theory of light-scattering spectrum of weakly bending filaments", *Macromolecules* **18**, 2430 (1985).

[25] K. Zimmermann, H. Hagedorn, C. C. Heuck, M. Hinrichsen, and H. Ludwig, "The ionic properties of the filamentous bacteriophages Pf1 and fd", *J. Biol. Chem.* **261**, 1653 (1986).

[26] J. Benjamin, B. K. Ganser-Pornillos, W. F. Tivol, W. I. Sundquist, and G. J. Jensen, "Three-dimensional Structure of HIV-1 Virus-like Particles by Electron Cryotomography", *J. Mol. Biol.* **346**, 577 (2005).

[27] R. Schwartz, P. W. Shor, P. E. Prevelige, and B. Berger, "Local rules simulation of the kinetics of virus capsid self-assembly", *Biophys. J.* **75**, 2626 (1998).

[28] M. F. Moody, “Geometry of phage head construction.”, *J. Mol. Biol.* **293**, 401 (1999).

[29] M. Piatak, M. S. Saag, L. C. Yang, S. J. Clark, J. C. Kappes, K. C. Luk, B. H. Hahn, G. M. Shaw, and J. D. Lifson, “High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR.”, *Science* **259**, 1749 (1993).

[30] R. L. Kingston, N. H. Olson, and V. M. Vogt, “The Organization of Mature Rous Sarcoma Virus as Studied by Cryoelectron Microscopy”, *J. Struct. Biol.* **136**, 67 (2001).

[31] D. M. Alamo and M. G. Mateu, “Electrostatic repulsion, compensatory mutations, and long-range non-additive effects at the dimerization interface of the HIV capsid protein.”, *J. Mol. Biol.* **345**, 893 (2005).

[32] G. B. Mortuza, L. F. Haire, A. Stevens, S. J. Smerdon, J. P. Stoye, and I. A. Taylor, “High-resolution structure of a retroviral capsid hexameric amino-terminal domain”, *Nature* **431**, 481 (2004).

[33] J. A. G. Briggs, M. N. Simon, I. Gross, H. G. Kräusslich, S. D. Fuller, V. M. Vogt, and M. C. Johnson, “The stoichiometry of Gag protein in HIV-1”, *Nat. Struct. Mol. Biol.* **11**, 672 (2004).

[34] J. A. G. Briggs, K. Grunewald, B. Glass, F. Forster, H.-G. Krausslich, and S. D. Fuller, “The Mechanism of HIV-1 Core Assembly: Insights from Three-Dimensional Reconstructions of Authentic Virions”, *Structure* **14**, 15 (2006).

CHAPTER 3

# NAVIGATING THE SEA OF CAPSID MORPHOLOGIES: EQUILIBRIUM CONCERNS

Now that we have looked at a simple growth model, we can begin to see some of the difficulties that arise in understanding the resulting closed capsids. Although equilibrium studies have shown that icosahedral capsids are energy minima [1, 2], our growth simulations rarely yielded icosahedral results. Moreover, it has been suggested that above $T = 7$, capsid assembly (without auxiliary proteins) is less reliable [3], possibly resulting in slightly non-icosahedral morphologies. This raises the question: what is the nature of the very lowest-energy excitations above the equilibrium morphologies? In particular, how do the elastic (excitation) energies of nearly-icosahedral capsids compare to $k_B T$? It appears that for large capsids, these low-energy excitations comprise a high near-degeneracy. In the following sections, we will explore this sea of capsid morphologies, including the near-icosahedral shapes, as well as the variety of tubes. We begin with an enumeration of all the capsid morphologies, followed by an exploration of the energy-minimizing structures as a function of our energy parameters. Finally, we discuss an attempt to compute the energy cost of small deviations from symmetry, as well as the issues arising from quantifying these deviations.

Throughout the chapter, we will continue to use the triangular network topology and Hamiltonian, as described in Chapter 2[1].

## 3.1 Enumeration of closed capsids

In order to discuss the ensemble of capsid morphologies, we need a way to talk about, and distinguish, individual capsids. Ideally, we would give every possible cap-

[1] Lidmar *et al* [4] used nearly the same model for their equilibrium work, though they refused to consider nonzero spontaneous curvature.

sid topology[2] a single name (preferably one that is simple and efficient to compute). However, coming up with a such naming scheme, by which identical capsids with different internal representations (i.e. reordering/relabeling the vertices) get the same unique name, is far from trivial. We denote the set of all capsids $\mathbb{Ca}$.

We make the simplifying assumption that each vertex in a well-formed capsid has coordination five or six[3]. Under these constraints, the Euler characteristic $\chi = 0$ requires that there be exactly twelve 5-coordinated vertices. Because the other vertices thus form a 6-coordinated background lattice, it stands to reason that the capsid morphology is completely described by the relationships between the 5-coordinated disclinations.

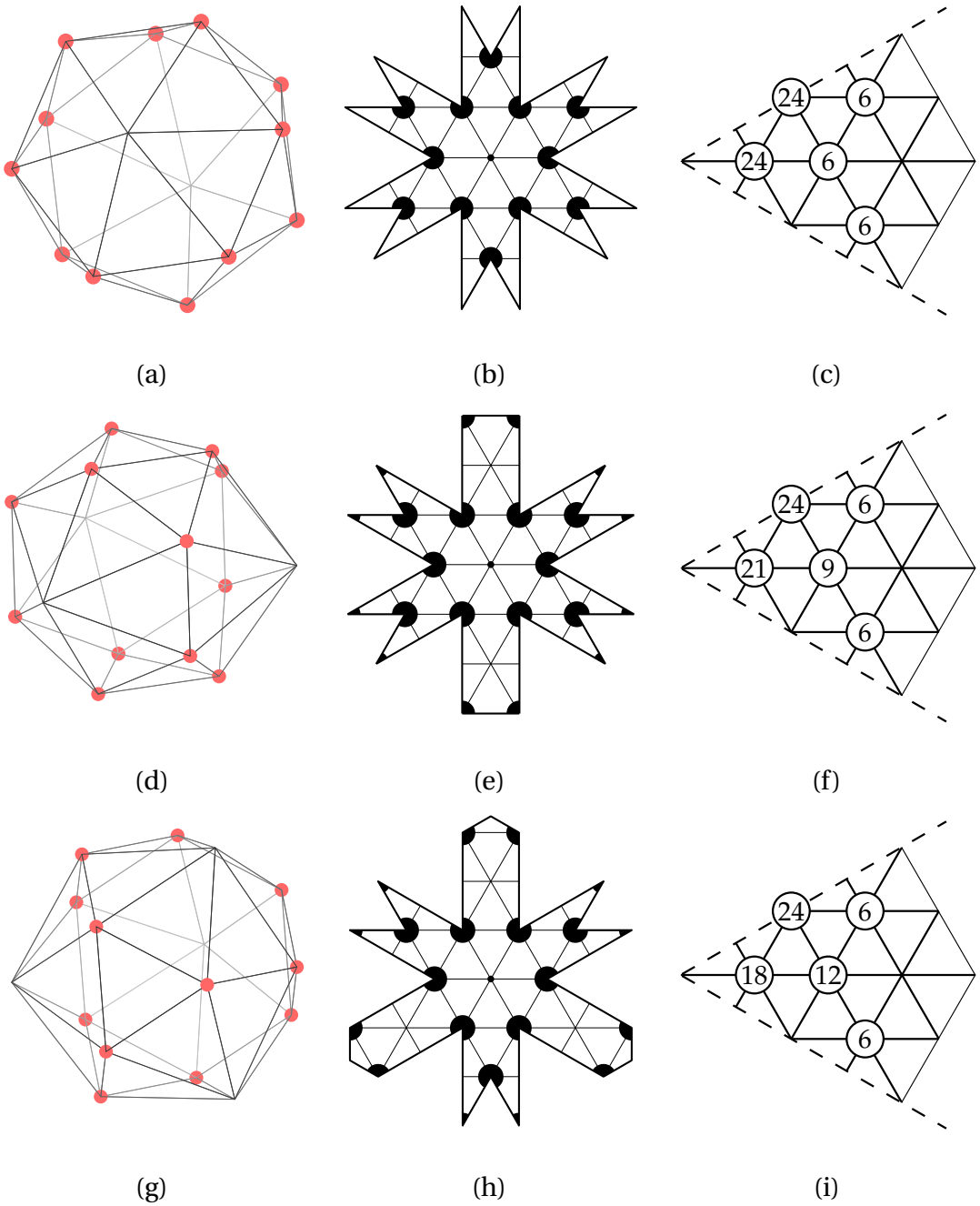
### 3.1.1 The Caspar–Klug framework

The most obvious naming scheme was given by Caspar and Klug [5] in the $T$-number system, where $T$ is the square of the lattice distance between every nearest-neighbor pair of disclinations in an icosahedral capsid. This gives a good start towards providing a unique name, but $T$ can be degenerate, either due to chirality or due to multiple pairs $(m, n)$ admitting integer solutions to $T = m^2 + mn + n^2$. Thus, we can get a unique label for icosahedral capsids by giving the specific $(m, n)$ decomposition, rather than just the $T$-number.

This $(m, n)$ decomposition can be understood as a follows: the distance $\sqrt{T}$ is traveled by first walking $m$ steps along a lattice direction, followed by a 60° turn and

[2]For our purposes, we define the *capsid topology* as the triangular polyhedral graph on the surface of a sphere, and independent of the real-space positions of the vertices. Thus, capsids differing by rotation, or even by a buckling transition, are considered identical, while chiral capsids differing by space inversion are considered distinct.

[3]Later sections in this chapter (§3.3.1) will allow intermediates to break this rule, but we are not interested in naming these intermediates.

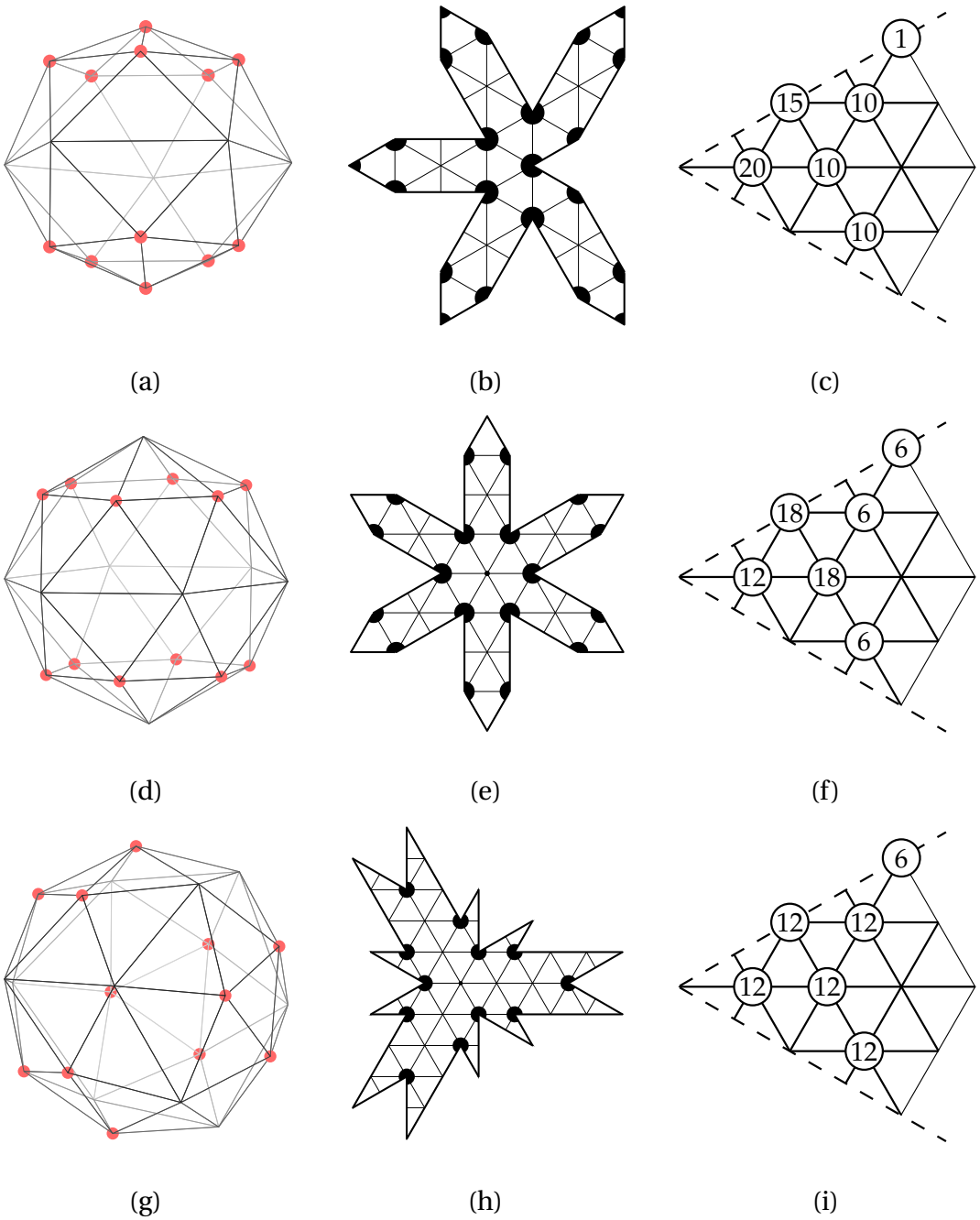



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 3.1: (a–c) The $T = 1.2$ capsid is a hexagonal antiprism with each hexagon replaced by six triangles. Alternately this is an $L = 0$ $(6,0)$ tube. (d–f) The $T = 1.3$ capsid consists of two large triangles (top and bottom halves of (d), or left and right halves of (e)) and three hexagons (around the waist of (d)). Alternately this is the shortest $(3,3)$ tube. (g–i) The $T = 1.4$ capsid is a truncated tetrahedron with each hexagon replaced by six triangles.




(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 3.2: (a–c) The $T = 1.5$ capsid is a length $L = 1$ $(5,0)$ tube, with pentagonal end caps. This begins a series of tubes with $T = 1 + L/2$. (d–f) The $T = 1.8$ capsid is a length $L = 1$ $(6,0)$ tube, with hexagonal end caps. This begins a series of tubes with $T = (6+3L)/5$. (g–i) The $T = 2$ capsid is a tetrahedron with faces each made up of ten triangles and three disclinations. There is an alternate (less symmetric) $T = 2$ capsid, which is the $L = 2$ $(5,0)$ tube, not depicted.

$n$ steps along the new direction[4]. Thus, for an icosahedral capsid, $(m,n)$ specifies the lattice displacement between a pair of pentameric vertices.

This notation is useful for tubes as well, by taking $(m,n)$ as the lattice displacement between a lattice point and its image around the circumference.

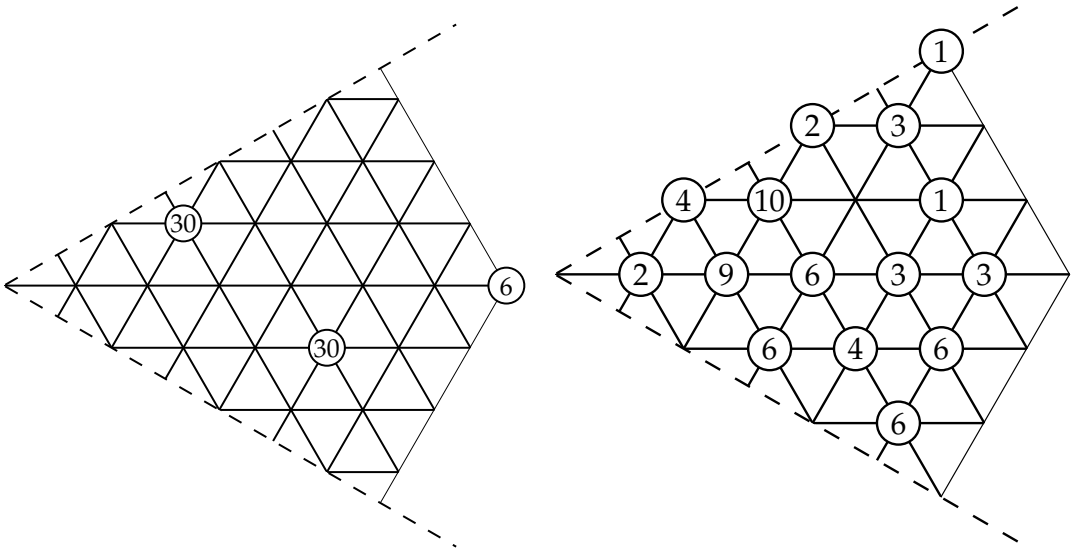
We can extend this framework slightly to describe some not-quite icosahedrally symmetric capsids as well. Figures 3.1 and 3.2 show all the capsids with between $N = 20$ and 40 triangles that have at least two different symmetry axes (another option would be to require at most two non-equivalent disclination environments). We can name these capsids uniquely with fractional $T$-numbers, consistent with $N = 20T$, as shown in the figure. At $T \geq 2$, however, we lose this uniqueness: there are two different capsids that could be called $T = 2$, and three different capsids that could be called $T = 2.2$.

### 3.1.2 Labels and representations

If we give up our requirement that names be bijective (that is, by allowing multiple capsids to have the same name, or single capsids to have multiple names), we end up instead with *representations* and *labels*. We will define a representation as a set $R$ together with a single-valued function $r : R \to \mathbb{Ca}$. On the other hand, a label is a set $L$ together with a single-valued function $\ell : \mathbb{Ca} \to L$. Obviously, a name is a function that is both a representation and a label.

The first, and simplest, label is to simply count vertices $V$, edges $E$, and faces

[4]It is convenient to interpret $n < 0$ differently in different cases: either as an opposite direction for the 60° turn and forward motion along the second direction, which is useful for describing chirality of capsids, or else as a consistent direction of turn and a negative walk in the new direction, which is more useful when looking at flattened capsid lattices; in this thesis we take the former interpretation, addressing the latter in §3.1.2.3




(a) (b)

Figure 3.3: Diagram of the total correlation function, showing the multiplicity of each lattice displacements between pairs of disclinations. The left-hand corner is the origin, representing the zero-displacement (we could add as "12" to this corner to represent the self-correlations if we wanted). The total of all numbers must always add to $66 = 12 \cdot 11/2$. (a) A perfect $T = 7l$ capsid. The 30's come from the five nearest and next-nearest neighbors of each disclination, while there are only six antipodal pairs of disclinations; (b) an irregular capsid with 92 triangles.

$F$. The Euler characteristic already constrains $V - E + F = 2$ for spherical topology. Moreover, our requirement that capsids have exactly 12 five-fold vertices allows us to calculate $E = \frac{1}{2}(12 \cdot 5 + (V-12) \cdot 6) = 3(V-2)$.[5] Thus, only one of $V$, $E$, and $F$ is actually independent, and this provides us with nothing more than a measure of capsid size.

#### 3.1.2.1 Hash labels

Another type of label (although this is actually more like a hash coding) comes from constructing an adjacency matrix between the vertices of the capsid. For a capsid with vertices $i = 1 \ldots N$, we define the $N \times N$ matrix with $A_{ij} = 1$ if vertices $i$ and $j$ share an edge, and $A_{ij} = 0$ otherwise. We then need an operation that is invariant

[5] We can extend this to the case where we additional 7–5 dislocation pairs: with $12 + D$ five-fold, $12 - 2D$ six-fold, and $D$ seven-fold, we still find $E = 3(V-2)$. If we allow $C$ four-fold vertices then we find $E = 3(V - C - 2)$.

under permutations of the vertex labels, such as the spectrum of eigenvalues, which reduces an $N$-vertex capsid to a list of $N$ real numbers in an almost-unique way[6]. Equivalently we could use the coefficients of the characteristic polynomial, or the traces of the first $N$ powers of the adjacency matrix, giving a list of $N$ integers related to the number of distinct loops of each length in the capsid lattice. Both of these methods provide rather long labels that are difficult to interpret, and even more difficult to reverse (i.e. construct a capsid with a given label).

#### 3.1.2.2 Correlation functions

As mentioned earlier, a capsid can be fully described in terms of the relationships between the fivefold disclinations. Thus, the most intuitive topological labeling is to list all the topological (lattice) distances (optionally retaining the $(m,n)$ coordinates) between every pair of disclinations (i.e. as a $12 \times 12$ matrix). This is naïvely a one-to-many approach, since reordering the disclinations also rearranges the distances (also, some distances are ambiguous, but can of course be recorded as such), but we can potentially reduce this to a single label for each capsid by deciding on a sorting rule.

If we flatten this matrix into simply a list of 66 displacements (ignoring the diagonal and the symmetric reflection) and count the number of occurrences of each displacement (equating modulo fivefold rotations, but not reflections), we end up with a total correlation function, which can be graphed on a 60° section of triangular lat-

[6] If two adjacency matrices $A$ and $A'$ have the same eigenvalues then there exists an orthogonal matrix $R$ such that $A' = RA$. The condition that $A$ and $A'$ be adjacency matrices with between five and seven neighbors for each vertex means that, aside from reordering of labels, it is rather difficult to construct such an orthogonal transformation, although it may not be impossible [6].

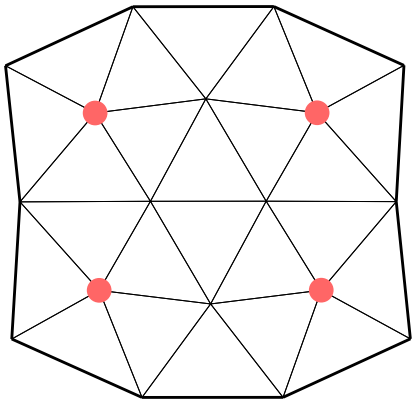


Figure 3.4: Degeneracy in Voronoi construction for nearest-neighbor disclinations.

tice, as shown in Figures 3.1–3.3[7]. Because these counts are discrete, we represent the function values by writing integers[8] at each lattice point in the wedge. This provides a nice visualizable label that gives a very clear picture of how regular a given capsid is. And while it's (most likely) not a representation, in that several capsids could have the same correlation function, it does appear to come very close (although it would be computationally difficult to reconstruct a capsid from a correlation function).

One could imagine also measuring something more like a direct correlation function by looking at only the nearest-neighbor disclinations as defined by a Voronoi construction. This provides a useful near-labeling of the capsid that gives a very clear depiction of the topology, except that the Voronoi construction is degenerate whenever four disclinations form the vertices of a rectangle, as shown in Figure 3.4.

[7] Heymann et al. [7] calls something similar to this a *Goldberg diagram*, but gives no primary source for that name.

[8] Some of the farther-separated disclination pairs do not have a unique shortest path since, e.g., $(5,1)$ and $(5,-1)$ are the same length. In cases where two such paths exist, it makes sense to count each possible shortest path with a fractional count, preserving the many-to-one nature of the graph.

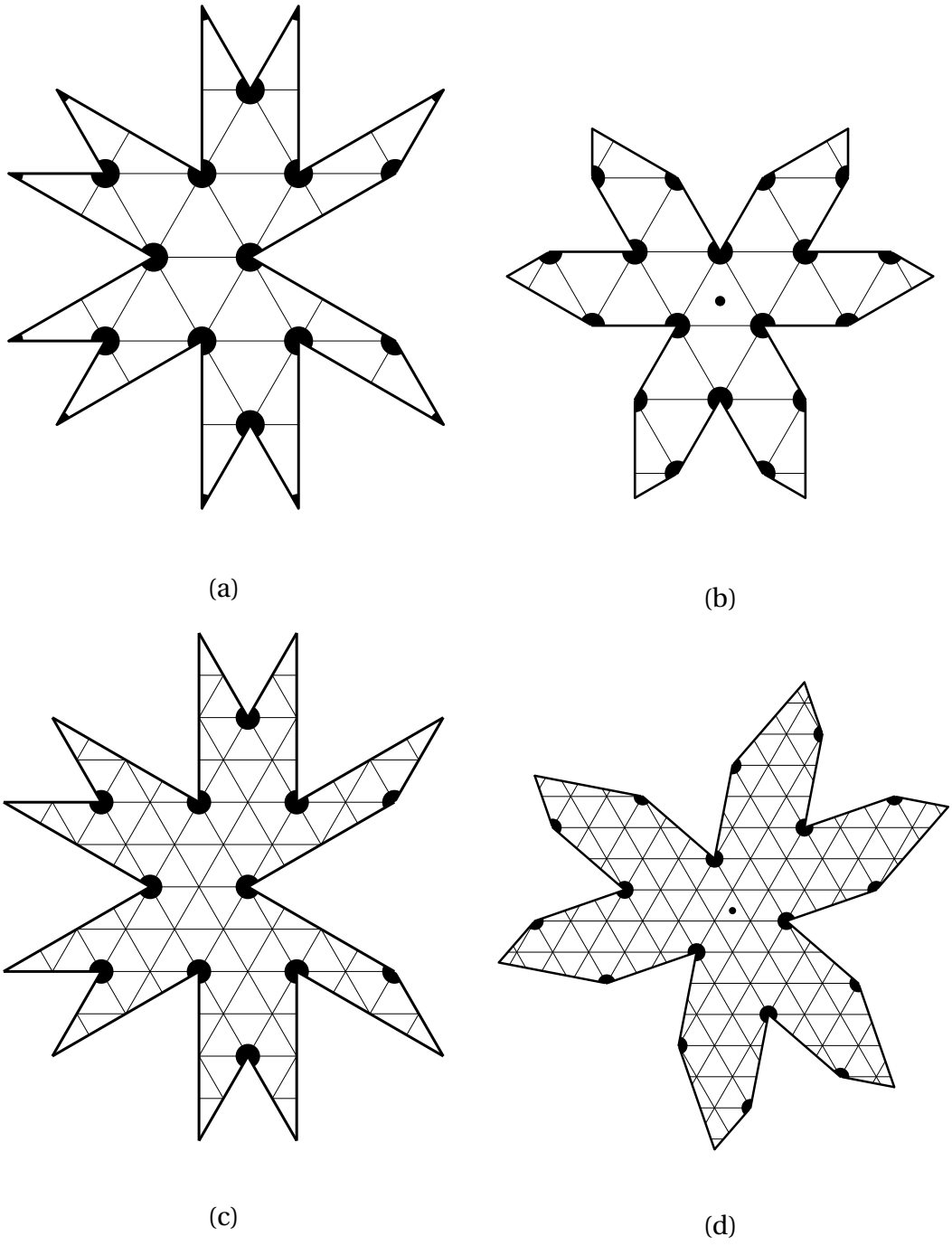



Figure 3.5: (a–b) Two flattened representations of a $T = 1$ capsid: in (a), one of the disclinations is chosen as the origin; in (b), the origin is chosen at a flat point. (c) A $T = 4$ capsid. (d) A $T = 7$ capsid.

#### 3.1.2.3 Flattened representations

Where the correlation function fails, the *flattened representation* succeeds, and vice versa: the correlation function provides an easy-to-compute label, while the flattened representation allows for easy reconstruction of a capsid from the representation. We can construct a flattened representation of a capsid by picking an origin and then cutting and flattening the capsid (each disclination must therefore intersect at least one cut) in such a way that for every point (not just the vertices), no alternate cutting would allow it to be closer to the origin[9]. Another way to think of this is that we are wrapping a piece of flat triangular lattice around the surface of a capsid. Clearly to do this we need to cut out triangular sections at each disclination, reattaching the two sides.

If two capsids have the same flattened representation, they are clearly the same capsid (but not vice versa). Because, as mentioned earlier, the capsid is fully-defined by the location of the disclinations, we can now fully define a capsid by giving the locations of the twelve disclinations on the plane (although each capsid will have many such representations, as shown in Figure 3.5).As with the disclination displacement matrix, we could possibly promote the flattened representation to be a naming scheme by defining rules to decide which of the many representations is canonical.

One benefit of flattening is that it is now easy to label points $\boldsymbol{r} = m\hat{\boldsymbol{\imath}} + n\hat{\boldsymbol{\jmath}}$ where

[9]Mathematically, let $\mathscr{C}$ be the set of points on the surface of the capsid, and $\mathbb{R}^2 = \{(r,\theta)|r > 0, 0 \leq \theta < 2\pi\}$ be a plane. Define a function $f : \mathbb{R}^2 \to \mathscr{C}$ by choosing an origin $O \in \mathscr{C}$ and picking a direction $\theta = 0$ at the origin. Then $f(r,\theta)$ is the point arrived at by traveling a distance $r$ (in lattice coordinates, i.e., assuming each triangle is equilateral and flat) out from $O$ in the direction $\theta$. (If such a ray passes through a disclination, we say it travels halfway around, but the choice turns out to be irrelevant since no shortest paths pass through disclinations anyway.) Note that $f$ is onto, but not 1:1. We then construct $g : \mathscr{C} \to \mathbb{R}^2$ such that $f(g(P)) = P$ and $|g(f(r,\theta))|^2 \leq r^2$ for all $(r,\theta)$. The flattened representation is then the set $\{g(P)|P \in \mathscr{C}\}$.

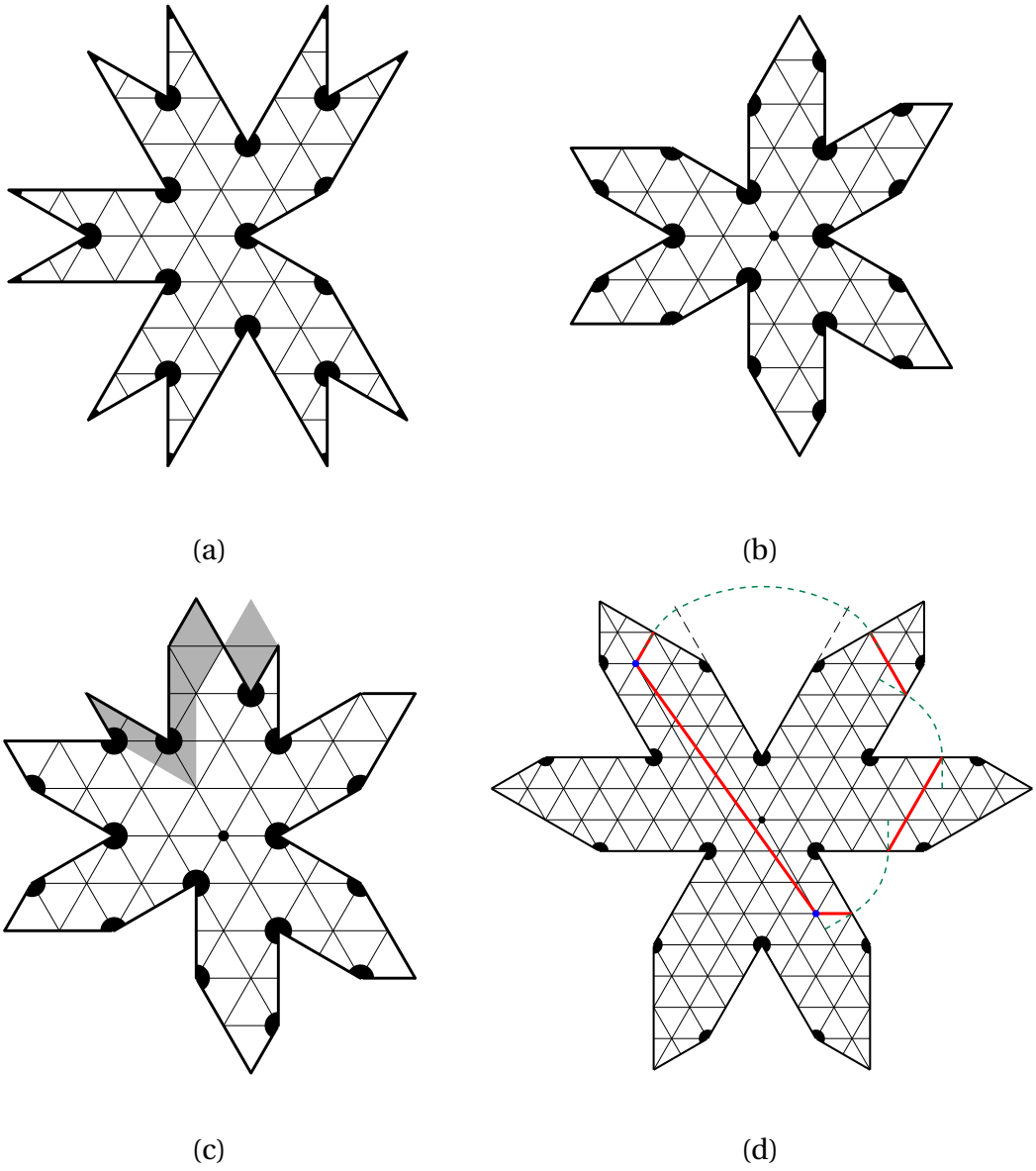



Figure 3.6: (a–b) $T = 3$ capsids with different origins chosen. (c) The result of moving two disclinations by one step each. Added and removed triangles are shown gray (there is a net gain of four). If we had not moved a pair of disclinations, we would have had a sevenfold vertex which cannot be represented in this fashion, except along a disclined edge. Note that this is not actually a proper representation: we would need to rotate the cut-out sections so that all points were in their closest directions. (d) Measuring the distance between two points on a flattened capsid is not trivial. In this case, the "direct" path has length $\sqrt{73} = \sqrt{8^2 + 8 \cdot 1 + 1^2}$, while the path that moves around the disclinations has length 7. The dotted arcs show the identification of the disclined edges, while the protruding dotted lines are perpendicular to the disclined edge, showing that the path remains straight.

$\hat{\boldsymbol{\imath}} = \hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{\jmath}} = \frac{1}{2}\hat{\boldsymbol{x}} + \frac{\sqrt{3}}{2}\hat{\boldsymbol{y}}$. [10] We also define $\hat{\boldsymbol{k}} = \hat{\boldsymbol{\jmath}} - \hat{\boldsymbol{\imath}}$ for convenience.

Another corollary is that we can more easily see the effect of moving disclinations. By determining the rules by which disclinations interact (topologically, not energetically), we can shift a disclination's position and propagate the effect on a flat medium, as shown in Figure 3.6(a–c).

While it is now trivial to measure the distance between any point and the origin (by definition), measuring other distances is not so straightforward, since crossing over a disclination section changes the angle of lines by $\frac{\pi}{3}$. Thus, the straight line connecting two points (a) may not be drawable without crossing a section, or (b) if it is, it may not be the shortest line, as seen in Figure 3.6(d). Thus, the best way to find the distance between two points is to reconstruct the flattened representation with the origin translated to one of the two points.

## 3.2 $T = 0$ phase diagram

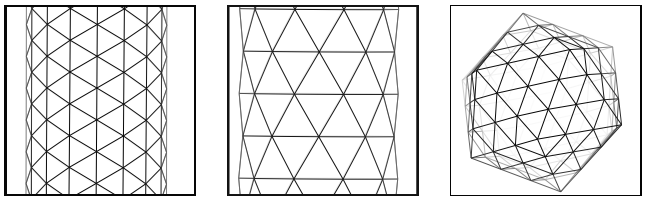
Now that we have discussed an enumeration of the set of capsids, the next step in exploring the low-energy excitations is to determine the energy minima. While Bruinsma et al. [1] showed that icosahedra (or tubes) are the minimum-energy structures, their phase diagram was of a relatively small part of their parameter space (which was of a qualitatively different kind than ours, since they parametrized on the hexon-penton switching energy), and they did not distinguish between different-sized tubes. We will therefore explore more details of different capsid morphologies in equilibrium.

[10] We could also look at it as a complex plane, with basis vectors of 1 and $\exp\left(\frac{i\pi}{3}\right)$, and then disclinations are somehow related to $z^{6/5}$ branch cuts.

### 3.2.1 Ensemble properties

The way our energies were written in Chapter 2, it is clear that a $T = 1$ capsid will generally have a lower energy than any larger capsid simply because it has fewer units. We might try to solve this problem by adding an edge-binding energy $E_b$, which is effectively equivalent to a chemical potential (see §4.1.5). We would thus be looking at a grand-canonical ensemble, in which we could tune the size of the capsid by changing the chemical potential of units. While this may be useful to consider, it does add an extra parameter to explore, so we will instead consider a canonical ensemble. Thus, suppose we have a gas of $M$ units. If we include them all into capsids with $N$ units each, and total elastic energy $E(N)$, then the total energy is $ME(N)/N$. Because (in the limit of large $M$) every unit is bound to exactly three other units, the total contribution of binding energy is $3ME_b/2$, which is fixed and therefore ignorable. Given a fixed $M$, we can see that for each set of elastic parameters, we must in fact minimize $E(N)/N$.

Since we are also considering tube geometries, we must mention end caps. While many tubular viruses do in fact have end caps, this introduces an extra parameter $L$: the length of the tubes. Fortunately, we can safely disregard this parameter: suppose there are $N_t$ units in the tube geometry and $N_s$ units in the sphere (end cap) geometry (we will ignore the interface between the tube and the end cap). If the tube geometry is truly preferred, then the minimum energy is at $L \to \infty$. Alternately, if the sphere geometry is preferred, the minimum is at $L \to 0$. Biologically we expect depletion effects or stochastic variation to limit the size of tubes, but in an equilibrium calculation, we can ignore $L$ and simply assume all tubes are infinite and thus have no end caps.



(a) (b) (c)

Figure 3.7: (a) a segment of (8,8) tube, (b) a segment of (10,0) tube, and (c) a (3,0), $T = 9$ "sphere", as defined in §3.1.1.

### 3.2.2 Inextensible case

If we reduce the problem to the inextensible case $Y \to \infty$, then we can calculate a few of the energies analytically—specifically, the $(m,0)$ and $(m,m)$ tube, and the $(m,0)$ icosahedron, as seen in Figure 3.7—since all the triangles must be equilateral. This is particularly useful in the case of tubes, in which we can completely eliminate the end effects that we necessarily see in numerical work (the other solution is to generate a single circumferential strip and use periodic boundary conditions in the axial direction, with a variable cell size).

First consider tubes. Due to symmetry, we need to consider only the bond angles on a single triangle. In the $(m,m)$ tube, as shown in Figure 3.7(a), it is clear that two of the three edges on each triangle have an angle of zero, since the tube is exactly an $m$-gonal prism. A quick calculation reveals the third angle to be $\pi/m$, so that the energy per triangle is

$$E_{\text{tube}}(m,m) = \frac{\kappa}{2\sqrt{3}}\left[\left(\frac{\pi}{m} - \theta_0\right)^2 + 2\theta_0^2\right]. \tag{3.1}$$

For the $(m,0)$ tube, we must take the $m \to \infty$ limit, so that we can make small-angle approximations. As seen in Figure 3.7(b), one angle on each triangle is negative (concave), and the other two are the same and positive. Under these assumptions, we

have angles of $-\frac{\pi}{m\sqrt{3}}$, $\frac{2\pi}{m\sqrt{3}}$, and $\frac{2\pi}{m\sqrt{3}}$, giving energy per triangle

$$E_{\text{tube}}(m,0) = \frac{\kappa}{2\sqrt{3}}\left[\left(\frac{\pi}{m\sqrt{3}}+\theta_0\right)^2 + 2\left(\frac{2\pi}{m\sqrt{3}}-\theta_0\right)^2\right]. \tag{3.2}$$

We can relate the tube radius $R$ to $m$ in each of these cases. For the $(m,m)$ tube, $R = \frac{m\sqrt{3}}{2\pi}$, and for the $(m,0)$ tube, $R = \frac{m}{2\pi}$ (taking $r_0 = 1$). If we write $m$ in terms of $R$ in each of these cases, then the energy

$$E_{\text{tube}}(R) = \frac{\kappa}{2\sqrt{3}}\left(\frac{3}{4R^2} - \frac{\sqrt{3}}{R}\theta_0^2 + 3\theta_0^2\right) \tag{3.3}$$

is the same for both $(m,m)$ and $(m,0)$, which makes sense since we might expect the bending energy to be isotropic. Given $\theta_0$, if we allow $m$ (or $R$) to be a continuous variable that minimizes the energy then we find

$$E_{\text{tube}}(\theta_0) = \frac{\kappa}{\sqrt{3}}\theta_0^2. \tag{3.4}$$

The $(m,0)$ icosahedron energy is even simpler. There are $20m^2$ triangles, with $30m^2$ edges, $30m$ of which are bent at an angle $\theta_{\text{icos}} = \arccos(\sqrt{5}/3)$, and the other $30m(m-1)$ of which are flat. Thus, the energy per triangle

$$E_{\text{icos}}(m,0) = \frac{\sqrt{3}}{2m}\kappa\left((\theta_0-\theta_{\text{icos}})^2 + (m-1)\,\theta_0^2\right). \tag{3.5}$$

For any given $\theta_0$, this is minimal for either $m = 1$ (a $T = 1$ icosahedron) or $m = \infty$ (a flat sheet). Then,

$$E(T=1) = \frac{\sqrt{3}}{2}\kappa\,(\theta_{\text{icos}}-\theta_0)^2, \tag{3.6}$$

$$E(\text{sheet}) = \frac{\sqrt{3}}{2}\kappa\theta_0^2. \tag{3.7}$$

These calculations are not particularly useful, however, since only considering $(m,0)$ icosahedra leaves all the $T$-numbers between $m^2$ and $(m+1)^2$, many of which are very energetically favorable. One simple observation we *can* make is that sheets are never favored over tubes at any nonzero $\theta_0$. To say any more, we will move on to a numerical study in order to access the more difficult in-between topologies.

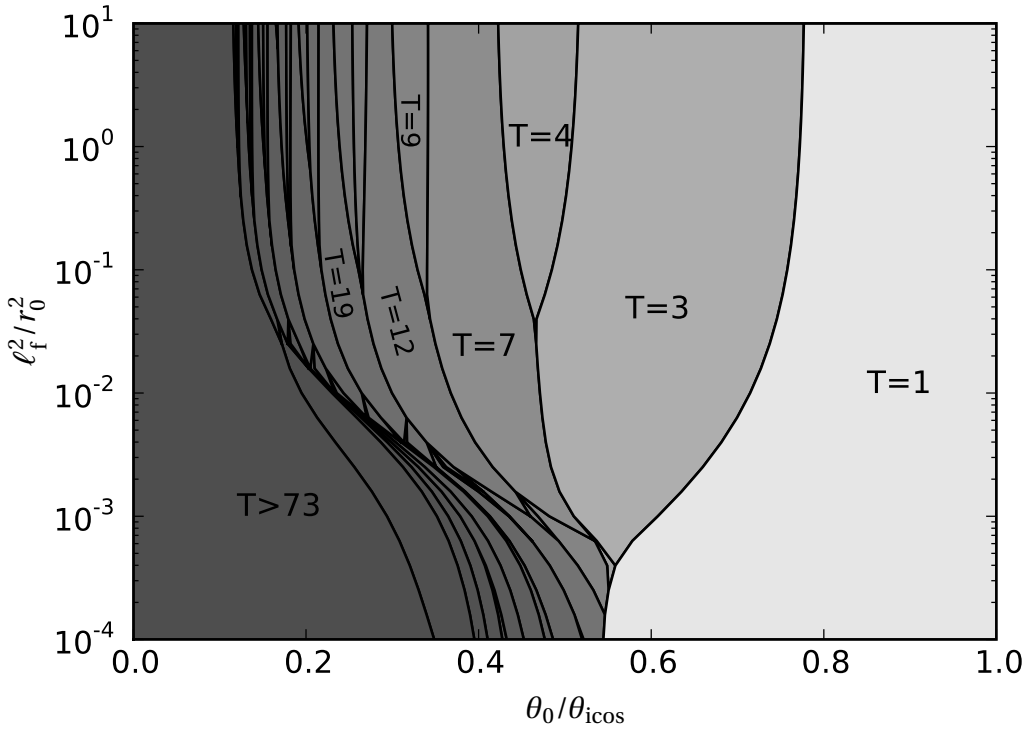



Figure 3.8: Phase diagram of icosahedral capsids as a function of $\theta_0$ and $\ell_f$. This diagram does not include tubes, and the vertical resolution of five values per decade causes visual artifacts.

### 3.2.3 Numerical results

In order to look at non-infinite stretching stiffnesses, we must resort to numerical computation. For complete capsids, changes in the spontaneous curvature $\theta_0$ have only marginal effect on the shape of the capsid. Thus, for a given capsid topology and a given value of $\ell_f$[11], we can efficiently fit the energy to a polynomial (primarily quadratic) function of $\theta_0$. Doing so for many different topologies allows us to calculate with sufficient accuracy the boundaries in $\theta_0$ between the "phases" in which a particular topology is preferred. Figure 3.8 is the result of computing these bound-

[11] In calculating the energies numerically, we find that capsid energies scale linearly in $\kappa/Y$ when $\kappa \gg r_0^2 Y$ or $\kappa \ll r_0^2 Y$, and are roughly constant throughout the buckling transition $\kappa \sim r_0^2 Y$

aries for many values of $\ell_f$ and interpolating lines between these rows. The larger values of $\ell_f/r_0$ (above about 1) are unphysical, since $\ell_f$ is comparable to the thickness of the capsid proteins (see §4.5.3).[12]

## 3.3 Low-energy excitations

Now that we know, for any choice of parameters $\ell_f$ and $\theta_0$, which capsid morphology minimizes the energy, the next step is to determine the minimum-energy deviations from these morphologies—i.e., the nearly-symmetric capsids. We can do this in a limited way by constructing a few specimens by hand, such as the fractional $T$-numbers in §3.1.1, or by attempting to make minimal adjustments to certain capsids by hand[13]. Beyond this, however, we would like an automated (e.g. Monte Carlo) approach. One such approach is to propose a transition from one capsid to a similar capsid, accepting or rejecting it based on the energy differences. We therefore need a method to generate transitions between closed capsids.

These transitions need a bit of cleverness. Simply removing a few triangles and reassembling is not suitable, because we need to move at least two disclinations in order to recover a well-formed (five- and six-fold vertices only) capsid (as we showed in §2.5, and will expand on in §3.3.1). Thus, we would need to melt back a rather significant amount in order actually change the topology. We will therefore consider alternate ways to generate complete-capsid transitions.

[12] This calculation was carried out in $\ell_f$-major fashion. That is, for each value of $\ell_f$, we picked five values of $\theta_0$ and minimized all the capsids for each set of parameters. Because the energy of a complete capsid is quadratic in $\theta_0$, we can use the quadratic fit of $E(\theta_0)/N$ to determine exact phase boundaries in the horizontal direction. The full two-dimensional phase diagram is then built up by stacking many of these one-dimensional phase diagrams on top of each other.

[13] For example, the minimal excitation from an $(m,1)$ icosahedron can be made by changing $m$ triangles.

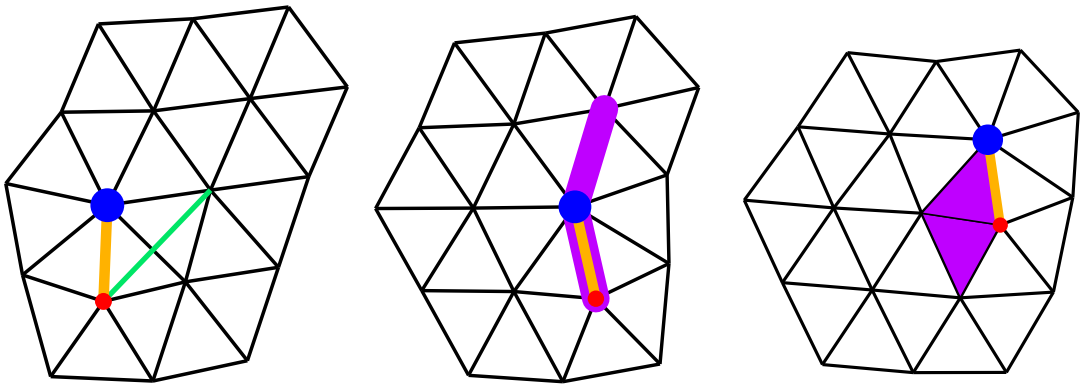


Figure 3.9: The dislocation moves. From left to right, flip and then climb; from right to left, descend and then flip. The dislocation is represented by the orange edge between a 5-fold and 7-fold vertex.

### 3.3.1 Dislocation movement

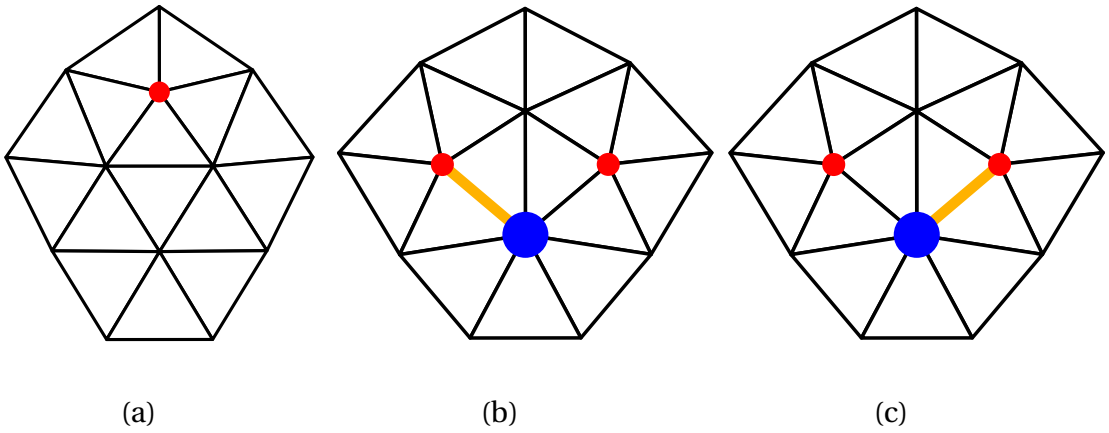
While we need to move disclinations in pairs to preserve well-formedness, we can get by with only local moves if we relax our constraints for intermediates by allowing sevenfold vertices, which form dislocations when paired with fivefold disclinations[14]. There are three elementary dislocation moves that can be applied to a complete capsid: *flip* (F), *climb* (C), and *descend* (D)[15].

The simplest transformation is the flip. Given two adjacent triangles, we have four edges forming a rhombus and the common edge is one of the two diagonals of this rhombus. We define a flip on the common edge as the transformation in which the one diagonal is replaced by the other diagonal, as shown in Figure 3.9(a↔b). This move is obviously self-inverting and conserves the number of triangles.

The other two moves, climb and descend, are inverses of one another. Climb (or create) starts with two connected edges in a line (i.e. sharing a common vertex but

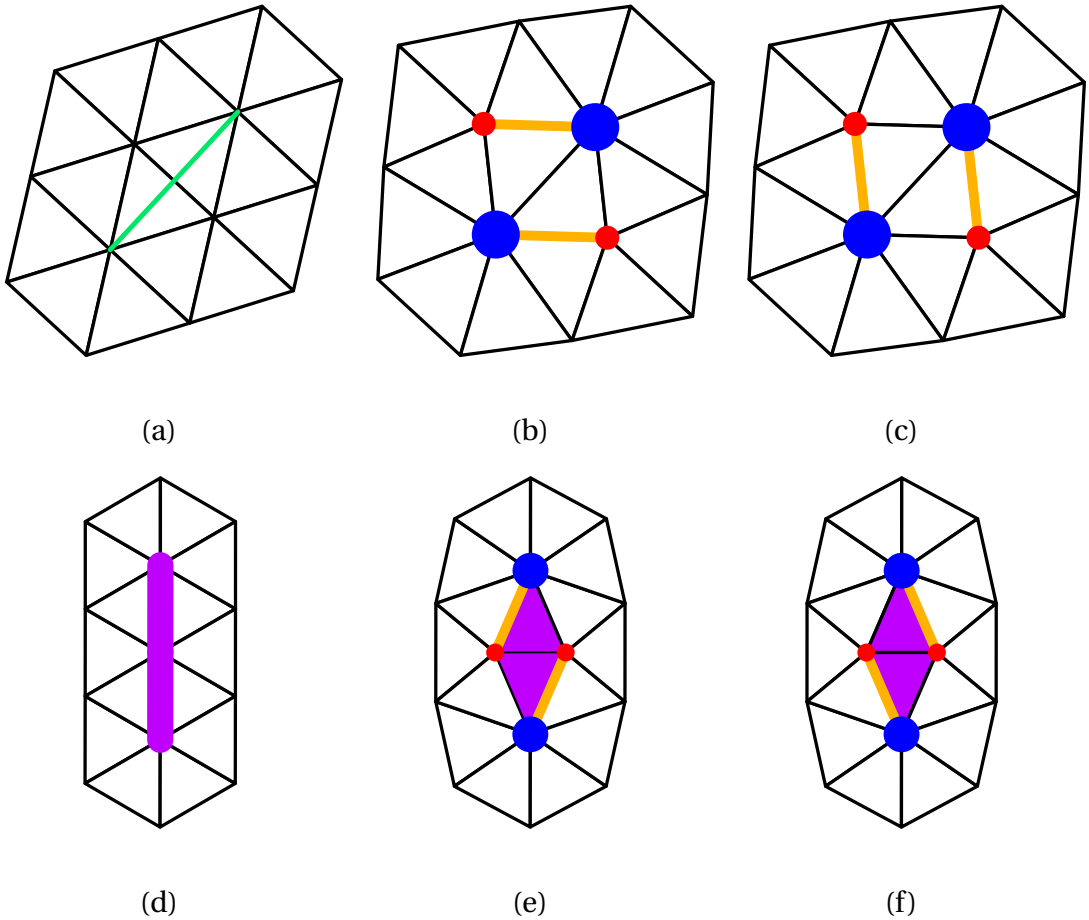
[14]While the 5–7 pair are typically adjacent, they need not be: a larger separation simply amounts to a higher dislocation charge.

[15]The latter two can also be remembered as *create* and *destroy*.



(a) (b) (c)

Figure 3.10: (a) Disclinations can be moved by one lattice step by creating a dislocation with a flip. The result can be interpreted in two different ways, (b) and (c), amounting to a choice in which direction the nucleated dislocation can travel.



(a) (b) (c)

(d) (e) (f)

Figure 3.11: Nucleating dislocation pairs with (a–c) a flip, or (d–f) a climb. Note that in each case, there are two different pairings that can be made.

not a common triangle) and cuts along the edges, inserting a pair of triangles in the hole, as shown in Figure 3.9(b→c). Thus, the two endpoint vertices each change degree by +1, while the middle vertex with degree $d$ splits into two vertices, with degrees $(d \pm \delta)/2 + 2$, with $\delta$ an integer. Climb increases the number of triangles by two. The inverse of climb is descend (or destroy). In this case, we find a pair of adjacent triangles and remove them, sealing the hole lengthwise, as shown in Figure 3.9(c→b).

Each of these moves can be used practically in three ways. First, these transformations move dislocations: flip is equivalent to dislocation glide, while climb and descend are technically both climbs. Second, using these moves, a dislocation can be created or annihilated by moving a disclination by one lattice space, as shown in Figure 3.10. Finally, we can apply it to a flat part of the capsid (i.e. with only six-fold vertices), effectively nucleating a pair of oppositely-directed dislocations, as shown in Figure 3.11.

Thus, with this vocabulary of transformations, we have effectively a recipe for generating new capsids that are "nearby" an existing capsid. These may be used as Monte Carlo steps in exploring the low-energy equilibrium ensemble of capsids, provided we can find steps that don't change the capsid too drastically, and that satisfy detailed balance. We will now present two such possible schemes.

#### 3.3.1.1 Flip-based update scheme

One possible recipe for generating Monte Carlo update steps is to pick any edge on the capsid and perform a flip to nucleate a pair of oppositely-pointed dislocations. We then repeatedly choose one disclination to glide, so that each choice is made with a $\frac{1}{2}$ probability, and each substep is fully reversible. We repeat these substeps until the dislocations annihilate one another. When a dislocation passes by a disclination,

we do have the option to deflect the dislocation by moving the disclination across the dislocation (as shown in Figure 3.10), but because dislocations are edge defects, while disclinations are vertex defects, we can simply disregard any interactions. Because everything is fully reversible, and because the capsid size remains fixed, detailed balance is automatically satisfied.

Unfortunately, this scheme suffers from the consequence that the moves are typically hundreds of steps long, causing drastic changes to the capsid. As the dislocations pass around different sides of the disclinations, their Burgers vectors become rotated with respect to one another, and their paths often cross (which is why we must be specific about which order we glide them in, since path crossing does not commute) before they can line up again to annihilate. The result is that we effectively generate a random element of the set of capsids of a given size. While this might be good in certain situations, our goal was to explore the space of nearly-symmetric capsids. Since these low-energy excitations occupy such a small fraction of the ensemble, taking such large steps means we are unlikely to ever return to this subset. We would thus prefer a "shorter" move that would allow us to remain in the low-energy well.

#### 3.3.1.2 Climb/glide update scheme

In an attempt to create a more local update scheme, it makes sense to pick a pair of disclinations and send a single dislocation from one to the other. This results in a minimal disturbance, and if we can find a way to do this reversibly, then we can still satisfy detailed balance.

Given any pair of disclinations $A$ and $B$, we can find the shortest path connecting them. We can then pick one of five directions to move $A$, nucleating a dislocation. In order to follow exactly along the shortest path to annihilate the dislocation

at $B$, there is a unique sequence of climbs and glides. Unfortunately, once a disclination has moved by a step, the shortest path may change (i.e. by jumping across another disclination). We instead use the shortest path from the shifted location of $A$ to the unshifted location of $B$. In this way, we have exactly $12 \cdot 11 \cdot 5$ different reversible moves, so each is chosen with the same probability.[16]

### 3.3.2 A metric over the space of capsids

Once we're able to generate near-symmetric capsids, we would like a way to quantify how "close" a given capsid is to being symmetric, or more generally, a metric over the space of capsids.

Given a set of growth rules (for instance, accretion and insertion from our irreversible model), we can define a tree-like graph in which each level contains (partial) capsids with all the same number of triangles. In this way, all edges are between nodes at levels $\ell$ and $\ell \pm 1$, and the complete capsids are terminal "leaves". (This graph is of course not a tree because it is not acyclic—there are many ways to reach the same result.) One theoretical metric, therefore, is to measure the minimal graph distance between two capsids. This would be roughly equal to (twice) the number of steps we'd need to backtrack in order to change from one capsid to the other.

A similar metric would involve populating the graph with capsids including dis-

[16] A final consideration is a few corner cases. What if the chosen disclinations happen to be located at the same place? Clearly we must allow four-fold vertices in order to maintain detailed balance (rejecting moves that result in fourfold vertices would unevenly increase other probabilities, without necessarily increasing the reverse probabilities), so we need to consider how to treat moves in their vicinity. On the other hand, these capsids are high-energy energy that they should be irrelevant to the exploration we're interested in. There's also the ambiguity of what to do when there are multiple equal-length paths. In this case, it should be reversible to choose one with equal probability.

locations (5–7 pairs, and possibly other coordinations as well, if several dislocations overlap), and adding lateral connections between complete capsids at the same or different level, by way of the *dislocation moves* explained in §3.3.1. This provides a sort of "Rubik's Cube" metric, in which we must solve the puzzle of finding the fastest way to transition from one capsid to the other given a set of elementary moves. Unfortunately, such a solution appears to be rather difficult, and we have not seriously attempted to solve it.

We can also employ real-space metrics of capsid distance. In general these rely on rotational alignment, which can be difficult because of disclination relabeling, but there are a finite number of relabelings to test in finding the optimal alignment. For any given labeling we can find the best rotation to align one capsid with the other (see §5.2.5), and then measure the real-space or angular differences in the location of the disclinations. Using purely angular distances cannot distinguish between any icosahedral capsids, even with different sizes, and degenerate icosahedra (i.e. $T = 49$, which can be broken up into $(7,0)$ or $(5,\pm 3)$) are indistinguishable by any measure that considers only the disclinations (of course, if the topologies of two capsids are different, then there is clearly no well-defined mapping between the six-fold vertices).

#### 3.3.2.1 Real-space characterization

We mentioned several measurements we carried out with our irreversible model in §2.3. These all have the disadvantage that they depend on the positions (and thus on elastic parameters), rather than only the topology.

There are a number of scalar measurements that produce only a single number, and thus are not particularly useful for classification. Most obvious are the average radius $\langle r \rangle$ (after translating the capsid so that the center of mass is at the origin) and

the asphericity, defined by $\langle (r - \langle r \rangle)^2 \rangle$ [4]. We already discussed in §2.3.3.1 an inverse participation ratio (IPR) that achieves a similar measurement of asphericity. Another measurement of facetedness (and this asphericity) is a histogram of the dihedral angles. In a faceted capsid, we expect a large spread in dihedral angle between neighboring units, while a smoothly spherical capsid will have a narrow distribution of dihedral angles. This is of course analogous to the radius-based measurements of asphericity, since the dihedral angle is directly related to the local radius of curvature, the average dihedral angle being inversely related to the average capsid radius. While measurements of asphericity are particularly useful for icosahedrally symmetric capsids, we are often more interested in assessing the amount of irregularity. In particular, we cannot distinguish an egg-shaped capsid from a faceted icosahedral on the grounds of asphericity alone.

We can, however, look at the covariances of the positions of all the vertices. In particular, if we define $\sigma_i^{(1)} = \langle r_i \rangle$, $\sigma_{ij}^{(2)} = \langle r_i r_j \rangle$, $\sigma_{ijk}^{(3)} = \langle r_i r_j r_k \rangle$, and so on, then we can look at the scalar invariants of the symmetric tensors $\sigma^{(n)}$. If the center of mass is at the origin then $\sigma^{(1)}$ is the zero vector. On the other hand, $\sigma^{(2)}$ is the moment of inertia tensor, which has three invariants (expressed in terms of the three matrix eigenvalues), and $\sigma^{(3)}$ has five invariants, which should describe the shape further, including a measure of egg-shapedness.

Moving beyond these scalar measurements, we consider the use of spherical harmonics for characterizing the shape of the capsid. There has been some work in the area of *icosahedral harmonics* [8, 9], which are particular combinations of the $\ell = 6j + 10k$ harmonics, for positive integers $j$ and $k$. In the case of non-icosahedrally symmetric capsids, we are also concerned with lower harmonics. In addition, there are a number of different choices of function to decompose into spherical harmon-

ics. We could look at the topological charge, which is simply the sum of twelve delta functions (one at each disclination), or the radius, or the Gaussian or mean curvature, or some measure of stress.

## BIBLIOGRAPHY


[1] R. F. Bruinsma, W. M. Gelbart, D. Reguera, J. Rudnick, and R. Zandi, "Viral Self-Assembly as a Thermodynamic Process", *Phys. Rev. Lett.* **90**, 248101 (2003).

[2] R. Zandi, D. Reguera, R. F. Bruinsma, W. M. Gelbart, and J. Rudnick, "Origin of icosahedral symmetry in viruses", *Proc. Nat. Acad. Sci. USA* **101**, 15556 (2004).

[3] J. E. Johnson and J. A. Speir, "Quasi-equivalent viruses: a paradigm for protein assemblies", *J. Mol. Biol.* **269**, 665 (1997).

[4] J. Lidmar, L. Mirny, and D. R. Nelson, "Virus shapes and buckling transitions in spherical shells", *Phys. Rev. E* **68**, 051910 (2003).

[5] D. L. D. Caspar and A. Klug, "Physical Principles In Construction Of Regular Viruses", *Cold Spring Harb. Symp. Quant. Biol.* **27**, 1 (1962).

[6] M. Kac, "Can one hear the shape of a drum?", *Am. Math. Monthly* **73**, 1 (1966).

[7] J. B. Heymann, C. Butan, D. C. Winkler, R. C. Craven, and A. C. Steven, "Irregular and semi-regular polyhedral models for Rous sarcoma virus cores", *Comput. Math. Methods Med.* **9**, 197 (2008).

[8] Y. Zheng, P. C. Doerschuk, and J. E. Johnson, "Determination of three-dimensional low-resolution viral structure from solution x-ray scattering data", *Biophys. J.* **69**, 619 (1995).

[9] Z. Yin, Y. Zheng, P. C. Doerschuk, P. Natarajan, and J. E. Johnson, "A statistical approach to computer processing of cryo-electron microscope images: virion classification and 3-D reconstruction.", *J. Struct. Biol.* **144**, 24 (2003).

CHAPTER 4

# FURTHER GROWTH MODELS

Now that we have more thoroughly explored the different morphologies of capsid shells, we will return again to the wide variety of growth models. We mentioned in Chapter 2 that assembly models are not universal, so we should expect to see different behavior from different model topologies.

We can categorize most growth models using three categories: the nature of the capsid representation, the nature of the free subunits, and the nature of the transitions between partial capsids.

The capsid representation has to do with the topology of the units in an assembled capsid, as well as the degrees of freedom and the functional form of the free energy. In all our previous discussion, we used a triangular lattice, loosely representing protein trimers, with degrees of freedom on the vertices and particular forms for the stretching, bending, and steric energies. In the following sections, we will consider the ramifications of other choices.

The treatment of free subunits is another important consideration. Our models so far have left them completely implicit, appearing out of nowhere at the growing edge of a capsid. We also assumed a canonical ensemble when comparing energies of different-sized capsids; but a grand canonical ensemble, with explicit free subunits, would have been reasonable as well.

Finally, we can define the transitions by specifying two things: which transitions are allowed (cf. the three growth steps in §2.1.3, though we might also want to include reverse steps; clearly this is coupled to the topology, but it also changes if we represent the free subunits explicitly), and when they occur (as a separate step, or simultaneous with the real-space dynamics).

We begin in §4.1 by exploring small perturbations to our model from Chapter 2: changes to the potentials or the growth rules, while maintaining the triangular lattice topology. In §4.2 we consider some more significant changes aimed at producing a model that grows the conical capsids characteristic of mature HIV. We then explore non-triangular topologies in §4.3, before introducing with a specific hexamer-based model for immature HIV assembly in §4.4. Finally, in §4.5 we fit the parameters of this hexamer model to experimental cryo-tomography data.

## 4.1 Triangular lattice models

Within the spectrum of triangular lattice models, there are still plenty of interesting changes we can make to our original model.

### 4.1.1 Nonharmonic potentials

The largest icosahedrally symmetric capsid our previous model was able to grow with any reliability was $T = 3$. One possible way to encourage the reliable growth of larger capsids is to use a bending potential with multiple wells. Since quasiequivalence requires $T > 1$ capsids to have several different local environments (which typically means different angles) we can think about a potential with local minima able to accommodate more angles without giving up the shape of the well in the immediate vicinity of the minima.

There are a number of less-obvious difficulties that arise in such an attempt. First, we already saw that the ratio of bending and stretching stiffnesses affects the shape of the capsid. In the inextensible, faceted ($\ell_f = \kappa/Y \to 0$) limit, an $(m,0)$ capsid has only two angles: $\theta_{\text{icos}} = \arccos(\sqrt{5}/3)$ and 0. At small $\ell_f$ (the smooth regime), we find instead $m$ different angles near $\theta_{\text{icos}}$ and $m(m-1)$ different angles near 0. As $\ell_f$

increases, these angles all tend towards a single limit $\theta_{T=m^2}$. Clearly with a multiple-well potential, $\ell_f$ is no longer well-defined, so the distribution of angles has already changed. We would therefore need to find a self-consistent solution (which isn't too hard, since the solution is iteratively stable). This stability unfortunately leads to another difficulty: because most capsid topologies have a somewhat broad range of angles, and are stable to changing these angles, we can realize a number of different-sized capsids with the same set of angles. In the extreme case, if we used a potential with wells at 0 and $\theta_{\text{icos}}$, we could realize any $T = m^2$ capsid with no strain. More generally, while there may be one targeted capsid that benefits greatly from the extra wells, nearby capsids will also benefit some, so that we don't gain as much discriminating power as we had hoped. One final difficulty is that we introduce a massive degeneracy in incomplete capsids, where each bond has multiple angles to explore. Our heuristic-based growth rules are no longer sufficient in such a situation.

### 4.1.2 Deterministic growth

Our irreversible growth model was entirely stochastic. Every possible transition had a non-zero probability, and so very unfavorable moves could be chosen, while very favorable ones were overlooked. Additionally, stochastic growth requires the growth to be repeated many times to get a representative sample. On the other hand, if the growth rules are deterministic, we need only grow a single capsid for any given set of parameters. This allows us much more quickly to explore the parameter space.

There are a number of ways we could implement a deterministic growth rule for our current model. The easiest is to retain all the same rates for each growth step, but to always make the move with the largest rate (this would correspond to the zero-temperature limit if our rates were based on energy differences). In our current

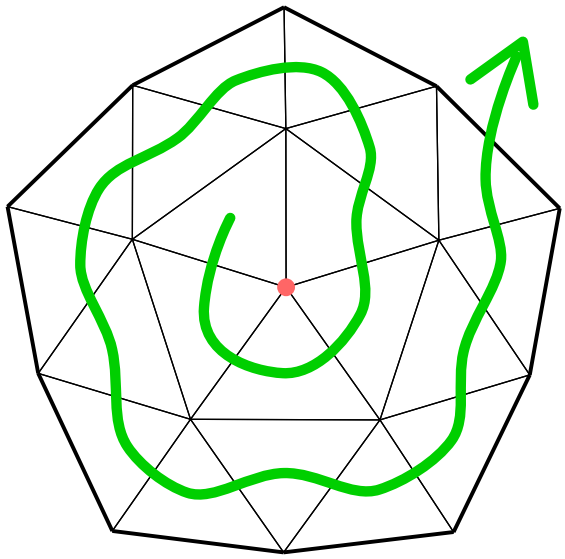


Figure 4.1: Deterministic growth spiraling outwards. When five triangles are around a vertex, the angle along the border is used to determine whether to insert a sixth triangle or to seal in a disclination.

model, when no vertices are eligible for insertion/joining there is a massive degeneracy because $\Gamma_A$ is constant. We therefore need to break the degeneracy for accretion by picking rates that depend on vertex coordination and angles. For instance[1], (1) if any vertex has five triangles about it, with an edge angle $\delta\alpha \equiv \alpha - \pi/6$ such that $|\delta\alpha| > \alpha_0$, then we either insert (+) or join (-) the vertex with the largest $|\delta\alpha|$, depending on the sign; (2) otherwise, for each edge, accrete on the edge for which the sum of the two neighboring angles $\alpha_i + \alpha_{i+1}$ is minimal, provided such accretion would not create a sixth triangle about either vertex.

A similar model to this was used by Levandovsky and Zandi [1], with the addition of a merging step between topologically distant parts of the capsid lattice.

Another approach for deterministic growth is to start with a single penton and grow spirally outwards, as shown in Figure 4.1. At each step, we add triangles to the

[1]To relate this to our model from Chapter 2, if we take $\sigma_I = \sigma_J$ and $\Gamma_I = \Gamma_J$ then $\alpha_0 = \frac{\pi}{6} - \sqrt{2\sigma_{I,J}^2 \log(\Gamma_{I,J}/\Gamma_A)}$.

border until there are five around a border vertex, and then use the angle $\alpha$ at that vertex to decide whether to insert a sixth triangle or to seal in a disclination, by comparing it to a threshold angle $\alpha_{\mathrm{T}}$. While such binary decisions allow us to explore the parameter space more effectively, it does severely limit the possible growth paths, and we lose the ability to put off deciding on ambiguous angles (i.e. near $\alpha_{\mathrm{T}}$).

This approach is particularly appealing because we can systematically explore the parameter space as follows: Given values for $\ell_{\mathrm{f}}$ and $\theta_*$, and a target capsid geometry (say, $T = 3$), we know ahead of time the required coordination at each vertex. Thus, we can keep running upper and lower bounds on the value of $\alpha_{\mathrm{T}}$ required to achieve that particular assembly. In particular, we can use the overlap (i.e. $\alpha_{\mathrm{T,min}} - \alpha_{\mathrm{T,max}}$) as a minimization objective function, varying the parameters $\ell_{\mathrm{f}}$ and $\theta_*$ in such a way as to decrease $\alpha_{\mathrm{T,min}} - \alpha_{\mathrm{T,max}}$ to the point of admitting the desired growth if $\alpha_{\mathrm{T,min}} < \alpha_{\mathrm{T,max}}$. Upon doing such a search, we find many parameters admitting $T = 1$ and $T = 3$ growth, a very small region admitting $T = 4$ (around $\theta_* \approx \pi/10$ and $\ell_{\mathrm{f}}/r_0 \approx 2$), while anything bigger turns out to be impossible under this model.

Another version of the spiral growth is that whenever we're confronted with a choice between two transitions (i.e. making a penton or a hexon), we can fully relax both possible choices and select the lower-energy result. Because of the small number of options at any given point, this is perfectly feasible, although it becomes more difficult to search for parameters producing a given result as we did in the previous paragraph with $\alpha_{\mathrm{T}}$, since we're only left with energy parameters. One simplification is to hold $\ell_{\mathrm{f}}$ fixed for a given target capsid, and then vary $\theta_*$ to find the crossover between penton and hexon formation to put lower and upper bounds on $\theta_*$. This adds quite a bit of computation time, but would be an interesting result.

### 4.1.3 Langevin growth

Every model we've discussed so far has had a clear separation between the "equilibrium" relaxation and the growth steps. We can, however, combine these two parts by allowing the positions to undergo stochastic fluctuations via Langevin dynamics, and continually monitoring the growing edge for possible growth moves (such as joins, in the case of the triangular lattice). Addition of new units would still need to happen with rates at each eligible edge, but these rates could depend on the local geometry as well. We effectively have three growth parameters: the rate of accretion $\Gamma_A$, the damping force $b$, and the threshold for joining $\alpha_J$; and we expect the most stable growth when $\Gamma_A$ is small. We have two different options for dealing with edges that never get close enough to join. First, we could leave the edges eligible for accretion, and when such an accretion occurs, we immediately join the new edge (effectively making it an insertion from Chapter 2), since the interior vertex now has six triangles and can no longer do anything else. This amounts to a race condition between the accretion and the fluctuation (how long it takes to fluctuate to a temporarily closed configuration to allow joining). Alternately, we could allow insertion within a certain threshold of $\alpha = \frac{\pi}{3}$, and disallow accretion near five-coordinated vertices, removing this race condition.

The amount of fluctuation of edge angles $\alpha$ is strongly dependent on the stretching stiffness of the capsid lattice. For particularly strong lattices (such as bacteriophage, with a three-dimensional bulk modulus of 1.4GPa) we find a root-mean-square deviation[2] $\sigma_\alpha \approx 1^\circ$. Many viruses have lattices that are much less stiff, so we

[2]We measure this by taking a large sample of relaxed partially-constructed capsids, picking a random vertex, and performing constraining relaxations, fixing the edge angle $\alpha$ to different angles, to determine $E(\alpha) \propto (\alpha - \alpha_0)^2$. The second derivative $\mathrm{d}E/\mathrm{d}\alpha$ together with the partition theorem gives the RMS fluctuations.

should expect fluctuations up to 5° in these capsids.

### 4.1.4 Reversible growth

After we completed the work described in Chapter 2, Philip Kidd spent some time exploring the possibility of adding backwards transitions to make the growth process reversible. This would allow for errors and even asymmetries to be annealed out of the capsid, thus allowing the capsid to potentially find a true energy minimum.

The basic approach was to add backwards versions of the growth rules $\overline{\mathrm{A}}$, $\overline{\mathrm{I}}$, and $\overline{\mathrm{J}}$ in a Metropolis simulation. We need to somehow determine rates for these transitions. If we assume equilibrium and detailed balance, then we would need to equate the probabilities for choosing A from one capsid and $\overline{\mathrm{A}}$ from the same capsid after an accretion. An easy way to balance these probabilities involves increasing the overall Metropolis rejection rate, but this slows down the process.

The simulations along these lines resulted in two types of capsids, depending on the parameters. With large curvature and a forward bias in the rates, $T = 1$ capsids assembled. Any other parameters with a forward bias resulted in monsters (that is, capsids where something went horribly wrong—either by forming multiple concentric shells, or making sharp angles, similar to the failures shown in Figure 2.4), and without the forward bias, most capsids didn't even complete.

### 4.1.5 Grand canonical ensemble

One other important consideration in a Metropolis simulation is that we need to be able to compare the energies of capsids of different sizes. This applies to irreversible growth simulations as well, if we are using a Metropolis-style acceptance criterion. In §3.2.1, we looked at a canonical ensemble, but in our growth simulations, we

are effectively assuming a bath of units, that is, a grand canonical ensemble. We could therefore assign a chemical potential $\mu$, such that the free energy of a capsid $E = E_{\text{elastic}} + N\mu$, where $N$ is the number of triangles. We could also consider a binding energy $-E_{\text{b}}$ for each non-border edge.

With $N_{\text{b}}$ border edges, the binding energy contribution is $-E_{\text{b}}(\frac{3}{2}N - N_{\text{b}})$. This effectively shifts $\mu$ by $-\frac{3}{2}E_{\text{b}}$ and adds a line tension $\tau = E_{\text{b}}$. All three of our moves alter the border length in different ways: $\Delta N_{\text{b,A}} = +1$, $\Delta N_{\text{b,I}} = -1$, and $\Delta N_{\text{b,J}} = -2$, with the reverse moves having the opposite signs. Thus, we have two parameters to adjust the bias toward either growth or disassembly, as well as toward pentons[3]. A significant amount of experimentation would be required to find reasonable ranges for these parameters[4].

Alternately, we might try to find some microscopic justification for a particular choice of these rates, based on the physical binding energies of units [2], the concentrations in solution (and the likelihood of meeting in a close-enough orientation). The latter might be explored with simulations of growth from a melt[5].

[3]Note that the parameters $\Gamma_{\text{A,I,J}}$ from Chapter 2 essentially correspond to a subspace of values for $\mu$ and $E_{\text{b}}$.

[4]We can do a bit of reasoning about some values without actually simulating anything. For instance, if $\mu - E_{\text{b}} < 0$ then accretions are essentially free, which would result in large chances of accreting when we really wanted to join.

[5]While we never attempted such growth-from-melt simulations, they are rather popular among a number of groups [3, 4, 5], and constitute a whole extra qualitative type of simulation beyond those we describe in this chapter. These simulations have an advantage over ours in that their growth parameters are better motivated, although the price is much more expensive computation and more noise among the data, since there is no single point of growth to concentrate on. Many studies, in fact, must rely on path sampling techniques [6, 7] because nucleation is such a rare event.

## 4.2 Models of mature retrovirus capsids

As discussed earlier in Chapter 2, irreversible growth models do not in general produce icosahedrally symmetric capsids. We therefore turn our attention briefly to retroviruses, which are known to form irregular capsids as well. In particular, mature HIV has a characteristic cone shape, as discussed in §1.3.

There are a number of simple additions we can make to the triangular network model in an attempt to produce conical capsids. In particular, we would like to address the question of which end of the cone forms first: Briggs et al. [8] suggest the narrow end forms first, while Benjamin et al. [9] implicate the fat end as the nucleation site, and Levandovsky and Zandi [1] conclude from their simulations that growth starts with a curled sheet. One other possibility we considered is that maturation doesn't involve a complete disassembly and reassembly as most biologists believe, but instead occurs by tearing out chunks of capsid lattice.

This is one case of a more general question—for any irregular virus (not just retroviruses), can we tell where assembly started and finished after-the-fact, using only the geometry and topology? In the case of cones, the rate of penton formation (see §2.3.1) needs to vary with time: in forming the small end, four or five pentons must nucleate very nearby; in forming the large end, seven or eight pentons must be sparsely distributed; in the part in between, there are no pentons. So the explanation for how cones assemble is tantamount to explaining what changes the penton formation rate.

### 4.2.1 Nucleic acid

It is well-known that mature HIV capsids will not assemble under normal biological conditions without RNA present(in order to assemble without RNA, the salt concen-

tration must be increased by orders of magnitude). However, it is not clear how to best model the interaction between mature CA and nucleic acid, since the NC proteins have been cleaved off and complexed with the RNA. The most naïve model is to assume that this complex is a point particle, interacting nonspecifically with the capsid. If the nucleic acid is very near the capsid then we might expect penton formation in order to increase the amount of capsid that is close to the nucleic acid (this would obviously require energy-based growth rules, rather than the geometry-based heuristics we described in §2.1.3). It might then be possible to nucleate a small number of pentons (particularly if the spontaneous curvature is small enough so that the capsid doesn't form tightly around the RNA, so that it eventually grows away from the RNA and is no longer affected). After this, the nucleation rate might be random, or perhaps even zero (if there is no spontaneous curvature) with the large end forming due to another effect, such as interaction with the membrane.

Ultimately, however, this is an unrealistic model of RNA. We could have corrected this by modeling it as a polymer, rather than a point particle, but it's unlikely that this would give conical results.

### 4.2.2 Membrane interactions

The reader should refer to §1.3 for a reference on retrovirus lifecycle.

Another explanation for the small end forming first is that the growing cone runs into the membrane. This says nothing about what causes the small end to form (perhaps the nucleic acid, or some other templating mechanism), and is motivated by the pictures showing the large end of the mature capsid always up against the viral envelope, and with the same curvature. This could be simulated by growing a capsid in a spherical potential $V_{\mathrm{m}}(r > R) = \frac{1}{2}k_{\mathrm{m}}(r - R)^2$, to represent repulsion by the envelope

(possibly with a larger power so that the elastic energy can't push the capsid too far into the simulated membrane). We never followed up on this idea.

Another possible cause of penton formation, related to the membrane, is a depletion effect. While it's not clear whether any free capsid proteins are left after maturation [10, 11], it may be possible in either case that a local depletion, caused by the geometry of the capsid up against the membrane, contributes a chemical effect as well as the mechanical effect considered in the previous paragraph: a larger chemical potential for free units would bias growth toward penton formation. Moreover, the effect would be cooperative, since it is believed that monomers aggregate into dimers (hexamers/pentamers would be theoretically possible, too) in solution before assembling into the capsid, so a small change in concentration would allow a large change in behavior.

## 4.3 Topological considerations

In general, when we look at a capsid representation, we must identify the underlying lattice, as well as the elements of the lattice that the degrees of freedom are associated with (the vertices, edges, and/or faces)[6]. While any reasonable lattice representation is topologically equivalent to a triangular lattice, different representations give rise more naturally to different degrees of freedom and interactions between the subunits, some of which may be more biologically faithful than others.

Figure 4.2 depicts the asymmetric units (monomers) of a capsid network as trape-

[6]A number of models, including those by Bruinsma et al. [12] and Nguyen and III [13], don't have an underlying lattice at all, instead distinguishing hexons and pentons by size and allowing non-specific interactions, so that "pentons" may actually have six neighbors.

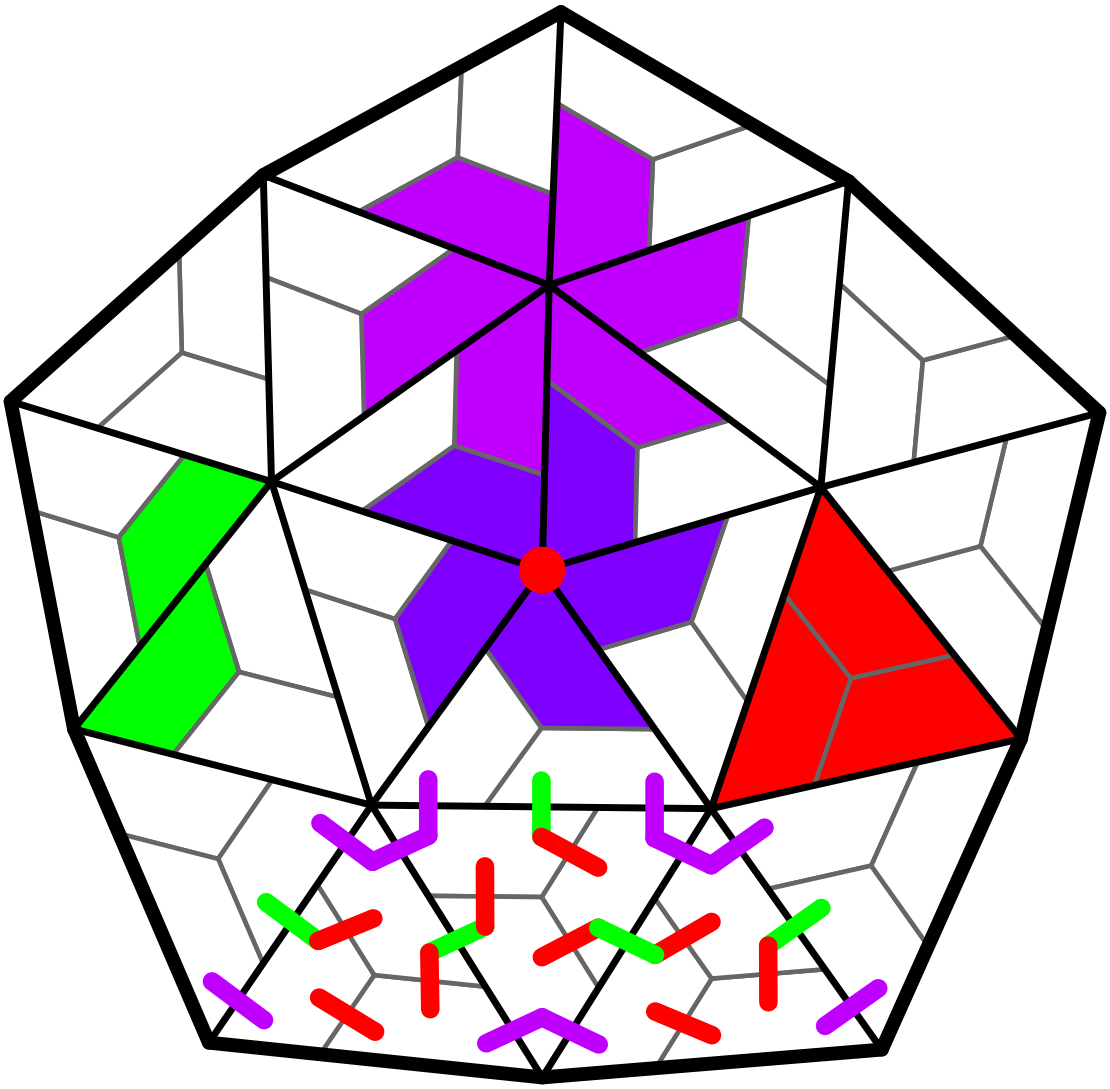


Figure 4.2: Cartoon of quasiequivalence in virus capsids. Trapezoids represent individual capsid proteins. All possible bond types are depicted with colored groups of units: a dimer in green, a trimer in red, and a hexamer (or pentamer) in blue (violet). The bonds are shown at the bottom in the corresponding colors. Of particular note is that two types of bonds are both necessary and sufficient to hold a capsid together.

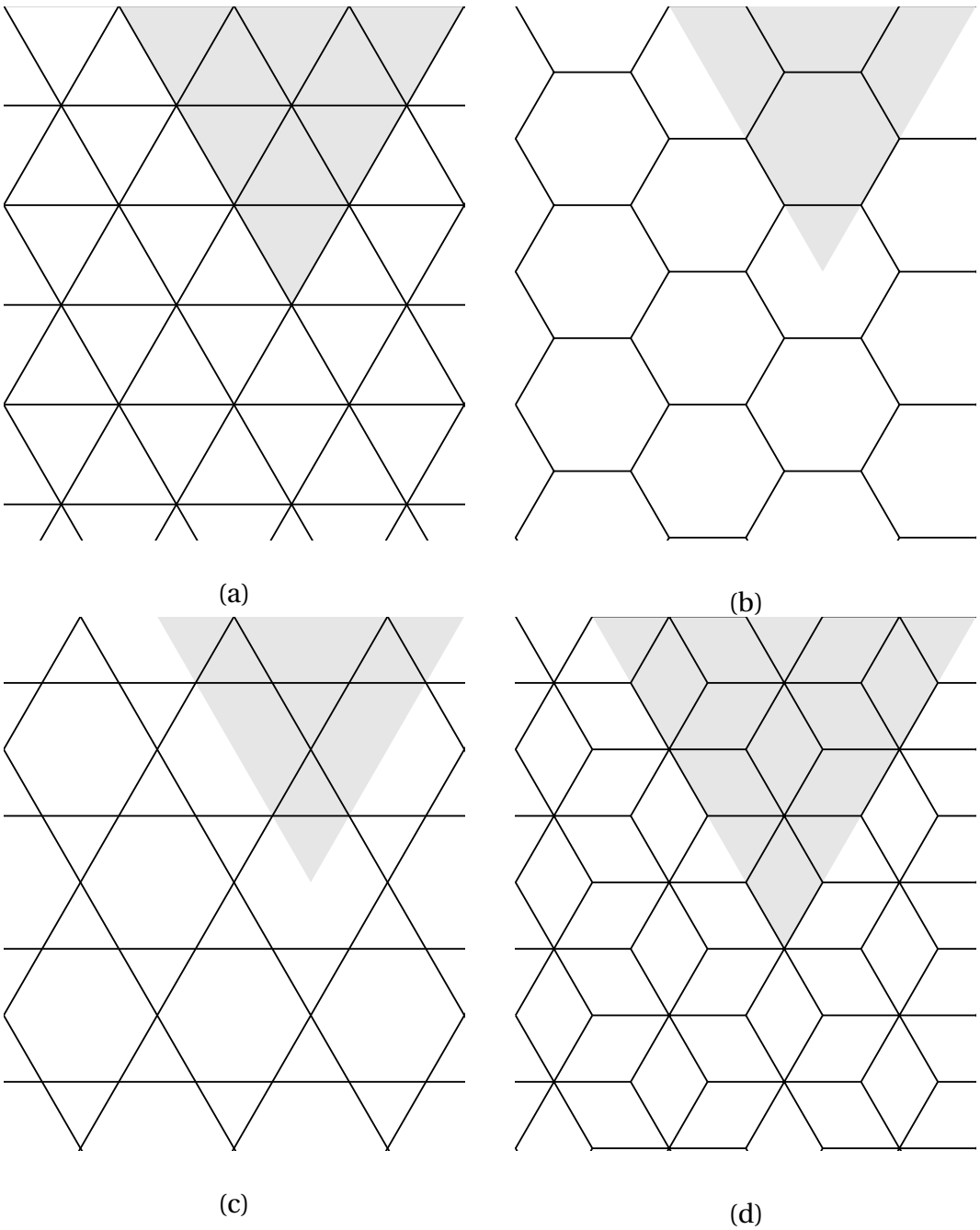



Figure 4.3: The four lattices allowed by symmetry: (a) triangular lattice, (b) honeycomb lattice, (c) kagome lattice, and (d) rhombille (or dice) lattice. The grey wedge in each lattice corresponds to the section of lattice that must be removed to represent a pentameric disclination.

zoids and shows all the possible types of interactions between them[7]. While we see three different types of interactions—dimer, trimer, and hexamer/pentamer—only two are required to build up a complete lattice[8]. There are four simple lattices with the correct symmetry, shown in Figure 4.3: triangular, honeycomb, kagome, and rhombille. In each lattice, the vertices, edges, and faces each correspond to dimers, trimers, or hexamers and pentamers. (It is worth noting here that we use "penton" and "pentamer", and likewise "hexon" and "hexamer", roughly interchangeably. A pentameric bond need not have come from a preassembled pentamer, for instance.)

In the triangular lattice, the triangular faces represent trimers, while the vertices represent hexamers and (for five-fold vertices) pentamers. Edges represent the dimer interaction. Thus, if we have degrees of freedom associated with each vertex, then our primary oligomer is a hexamer/pentamer. We can then choose to write an interaction energy as a function of the edges (in which the dimer bond is secondary to the hexamer bond), or else as a function of the triangles (i.e. a three-body interaction, in which the trimer bond is secondary to the hexamer bond). Alternately, we could attach degrees of freedom to each triangular face, so that the trimer is the primary unit, and then choose either a two-body edge-based interaction, or a 5- or 6-body vertex-based interaction. The honeycomb lattice is dual to the triangular lattice, so the vertices and faces exchange roles.

The kagome lattice is slightly more complicated, in that we have two different kind of faces. Note that we cannot take both hexamers and trimers as our primary

[7]Note that while this was motivated by standard CK quasiequivalence [14], it fits loosely into Twarock's tiling theory framework [15] as well by identifying rhombs with the dimers and kites with the trimers, and pulling out one trapezoid of each hexamer at the interface, stretching it into a pentamer.

[8]However, some monomer-based models have explicitly modeled all three [16], and it's perfectly possible for a single virus to exhibit all three types of bonds.

unit at the same time, since the two units have monomers in common. The vertices here represent dimers, and the edges represent the internal interactions in the primary unit. By associating our degrees of freedom with either type of face or with the vertices, we can take any of the three oligomers as the primary unit, and by constructing an interaction at either the vertices or around the other type of face, we can choose any secondary bond type. The rhombille (or dice) lattice is dual to the kagome lattice, and thus the vertices and faces exchange roles. Note that there are now two different types of vertices: threefold and sixfold (fivefold in the case of pentons), and these must be treated separately. This lattice therefore does not lend itself well to vertex-based representations.

In addition to these four simple lattices, we can construct hybrids of a lattice and its dual, which we will describe in §4.4.

### 4.3.1 Sources of flexibility

While we've so far not explicitly specified what exactly the attached degrees of freedom represent, there has been an implicit assumption that they are positions of rigid units. Because we need to have some sort of flexibility in our lattice, this requires us to assume that the bonds between units are flexible. We could consider the opposite case, however, in which the bonds are rigid and the units are flexible. An extreme example of this on a triangular lattice might be a model in which we require each pair of triangles to be attached smoothly at the edges, and to have no angle deficit (from $2\pi$) at the vertices, such that all the Gaussian curvature is contained within the triangles themselves, rather than in the edges and vertices. In such a model, the angle abundance in each triangular unit would be the source of Gaussian curvature, and

we could construct a Hamiltonian to favor a particular spontaneous curvature[9].

Our model in Chapter 2 was somewhere between these two extremes. We imagined (by our growth rules) that our units were represented in the triangular faces, which were flat but allowed to stretch or shrink. Thus, the stretching elasticity was contained in the units, and the bending elasticity in the bonds.

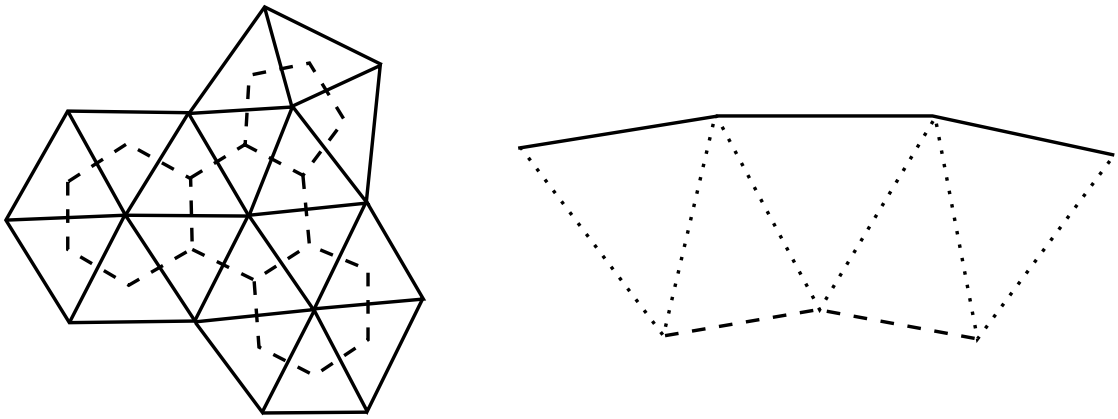
### 4.3.2 Universality

While these different pictures are largely redundant with respect to mechanical equilibrium, they have a significant influence on growth: in the simplest picture of growth, where a single capsid nucleates and grows toward completion, the transitions between intermediate structures depend fundamentally on the units out of which they are built. In our model of irreversible assembly of a triangular lattice, we found that the transitions required redundant growth rules (join and insert) in order to achieve successful growth (see §2.1.3.1), while a similar model made out of hexons and pentons would avoid some of the failure modes we ran into. Hagan and Chandler [5] investigated three different lattices and found different qualitative behavior.

### 4.3.3 Multi-layer network models

After working out our initial assembly model, we found a number of disadvantages to the triangular network model. Biologically, the assumption of trimeric units is rarely supported, with very few known cases of trimers in solution (one example is lambdoid phage 21 [17]). Secondly, the requirement to have both the join and insert rules

[9] Such a model might be useful in integrating out the positional degrees of freedom, if we could write matching rules relating the angles around the triangle, the edge lengths, and the bending curvature of each edge, since we would no longer need to worry about absolute positions.



(a) (b)

Figure 4.4: Two-layer network model. (a) top view; (b) side view. The solid network is triangular and is shown on the outside, while the dashed network is honeycomb and is shown on the inside, though it could just as easily be turned inside out with different spring lengths. The dotted lines in (b) show the springs connecting the vertices of each network. The triangles could represent either trimers or monomers; in the latter case, the center hexagon would not be a hexamer and therefore the three of its edges connecting the two hexamers and the pentamer could have a different spring constant.

was rather distasteful, and this is avoided in a hexon-based model, where a viable model should be possible with only addition-based rules, since the vertex topology is simpler (always three units around each interior vertex). Finally, because our units were flat, the steric potential was difficult to deal with, since we couldn't use simply a volume exclusion. The bending stiffness was also slightly artificial, based on triangle normals (this might be seen as an advantage, however, since it allowed the stretching and bending moduli to be varied fully independently from one another).

As an extension, therefore, we looked for a hexamer-based model. The most successful versions of this involved coupling a pair of two-dimensional spring lattices—an inside and an outside—by another set of springs. By making one lattice triangular and the other hexagonal (that is, a honeycomb lattice with pentagonal disclinations), as shown in Figure 4.4, we could control the spontaneous curvature and the Foppl-

von Kármán ratio by the choice of lengths and stiffnesses. We found an additional benefit of being able to impose effectively different interactions between pairs of hexamers as between a hexamer and a pentamer, since the degree of puckering in the pentamer is different. This distinction is fully consistent with quasiequivalence, and could be useful in indirectly simulating, for instance, a double-well angular potential (as discussed in §4.1.1) that might more easily produce $T = 4$ or even $T = 9$ capsids.

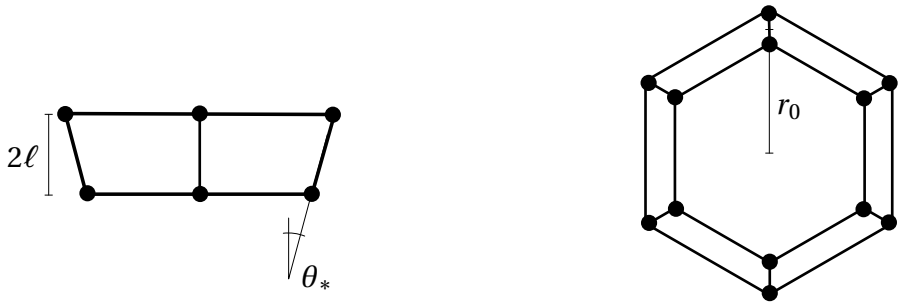
While Rapaport [16] and Levandovsky and Zandi [1] have three-dimensional units, there is still only one lattice topology. The distinction we're making here is that one can pair two lattice topologies by putting one topology on each layer.

On the other hand, this approach still ties the positional degrees of freedom to the lattice vertices, and as such there is no good way to favor a twist (see §5.2.6) between units, which might be helpful in selecting the chirality of $T = 7$ capsids.

## 4.4 Immature HIV assembly

Immature HIV is known to assemble on the plasma membrane of infected cells, simultaneously curving the membrane outwards (although budding does not actually take place without help from the cell). It has been shown, however, that immature capsids can assemble with the same morphologies in the absence of membrane, so we will ignore the effects of the membrane for the time being.

An important discovery with respect to immature HIV is that there appear to be holes in the lattice in exactly the places we would expect to find the pentameric disclinations [11, 18]. Thus, we need only consider hexameric units in our model. We will now present a model that differs rather significantly from those discussed above—in particular, because our degrees of freedom are no longer equivalent to the positions of lattice vertices, but rather represent the position and orientation of individual sub-




(a) (b)

Figure 4.5: Hexamer-based model for immature HIV assembly: (a) side view, (b) top view. The filled circles represent the twelve contact points on each unit.

units. This separation allows a more general interaction (cf. Chapter 6), since relative motions that were implicitly disallowed, such as twist, are now possible. In addition, the subunit-centric approach of assembly from melt as mentioned in §4.1.5 may be more viable under this model.

We start by specifying our hexameric units and the interactions between them in §4.4.1. We then discuss alternate variations that we considered previously in §4.4.2. Following this, we look briefly at the kind of growth rules we would need in order to effectively simulate immature HIV growth in this model in §4.4.3. We then go on to fit our model parameters with experimental data in §4.5.

### 4.4.1 Hexamer units

Our units are effectively hexagonal frustums. We parameterize each unit with a vector for the center $\boldsymbol{r}$ and an orientation $\boldsymbol{\Omega}$, which defines three orthonormal vectors $\hat{\boldsymbol{n}}$, $\hat{\boldsymbol{t}}$, and $\hat{\boldsymbol{b}}$, where $\hat{\boldsymbol{n}}$ is the outward-facing normal to the hexamer, and $\hat{\boldsymbol{t}}$ points from the center of the hexamer to a lateral edge (and thus $\hat{\boldsymbol{b}}$ points to a lateral face). Each unit then has twelve vertices, at

$$\boldsymbol{v}_{m\pm} = \boldsymbol{r} + (r_0 \mp \ell \sin\theta_*)\left(\hat{\boldsymbol{t}}\cos\frac{m\pi}{6} + \hat{\boldsymbol{b}}\sin\frac{m\pi}{6}\right) \pm \hat{\boldsymbol{n}}, \tag{4.1}$$

where $m = 1 \ldots 6$, $r_0$ is the circumradius of each unit, $\ell$ is the height, and $\theta_*$ is the spontaneous curvature. We can form a unitless parameter $\ell / r_0$ that relates to the Foppl-von Kármán ratio.

Each hexagonal unit has six trapezoidal lateral faces, upon which we define a bonding interaction. The interaction energy for each bond is given by placing four equal zero-length harmonic springs: one between each corresponding pair of vertices of the interacting faces, i.e.

$$\begin{aligned} E_{\nu m;\nu' m'} = \frac{k}{2}\Big( &\left(\boldsymbol{v}_{m+} - \boldsymbol{v}'_{(m'+1)+}\right)^2 + \left(\boldsymbol{v}_{(m+1)+} - \boldsymbol{v}'_{m'+}\right)^2 \\ &+ \left(\boldsymbol{v}_{m-} - \boldsymbol{v}'_{(m'+1)-}\right)^2 + \left(\boldsymbol{v}_{(m+1)-} - \boldsymbol{v}'_{m'-}\right)^2\Big). \end{aligned} \tag{4.2}$$

While this is the simplest possible model that meets our requirement of supporting a spontaneous curvature, a fully-harmonic interaction is only suitable for interactions between topologically neighboring units. In order to generalize to topologically distant units, we would need to pick a short-range function that vanishes past a certain cut-off length $x_0$. The simplest solution here is to use a cosine function,

$$\frac{1}{2} k x^2 \to k x_0^2 / \pi^2 \left(1 - \cos(x\pi / x_0)\right), \tag{4.3}$$

but this implicitly relates the stiffness to the binding energy by the cut-off length, which may be undesirable (although with no microscopic motivation, it may be as reasonable as any other choice).

Note that although we have four "contact points" on each trapezoidal binding edge, we only strictly needed three points for a rigid energy minimum, whereas with only two points we have an unconstrained bending degree of freedom.

We can compare our new parameters $k$, $\ell$, $r_0$, and $\theta_*$, to the parameters from Chapter 2, which we'll now call $Y$, $\kappa$, $r_\triangle$, and $\theta_\triangle$. Indeed, is is straightforward geometry to see that

$$r_\triangle = r_0 \sqrt{3}, \qquad \theta_* = \theta_\triangle \sqrt{3}. \tag{4.4}$$

The energy parameters are more difficult, but comparing the energy from each models results in

$$8k = Y\sqrt{3}, \qquad 4\ell_{\mathrm{f}}^2 = 3\ell^2, \tag{4.5}$$

where, recall, $\ell_{\mathrm{f}}^2 \equiv \kappa^2 / Y$.

### 4.4.2 Alternate formulations

We considered several other alternatives to this formulation. Initially, our hexagons were planar, so that the lateral faces in this model were replaced with lateral edges having only two interacting vertices. This interaction would take care of the stretching stiffness, as well as a stiffness against two of the three orientational deformations (all but bending). In order to achieve any stiffness to bending, we would need an interaction between neighboring units' normal vectors $\hat{\boldsymbol{n}}$. The $1 - \hat{\boldsymbol{n}}_1 \cdot \hat{\boldsymbol{n}}_2$ term we used in the triangular network model would work only in the case of zero spontaneous curvature. Our modification for the triangular network, using $(\hat{\boldsymbol{n}}_1 \times \hat{\boldsymbol{n}}_2) \cdot \hat{\boldsymbol{e}}$ to calculate the sine of the angle, no longer works in this situation because the common edge $\hat{\boldsymbol{e}}$ is no longer well-defined, since the units are rigid. We could attempt to reconstruct it by averaging the two binding edges $\hat{\boldsymbol{e}}_i$, or by averaging the two $\hat{\boldsymbol{n}}$ vectors, crossed with the displacement between the two hexamers' centers $\boldsymbol{r}_{12}$, or possibly make use of $\hat{\boldsymbol{n}}_1 - \hat{\boldsymbol{n}}_2$, but all these approaches are equally problematic: because twist is now allowed, we have an extra dimension in our bending degree of freedom, so that the task of writing a potential between the unit vectors can be reduced to specifying a function over the 2-sphere.

In order to assure a minimum at a specified angle $\theta_*$ from $\hat{\boldsymbol{n}}_1$, there are two immediately obvious options. First, we could simply rotate $\hat{\boldsymbol{n}}_1$ by $\theta_*$ through an axis constructed from $\hat{\boldsymbol{e}}_i$ or $\boldsymbol{r}_{12}$. As mentioned above, there is not a single well-defined

way to construct this axis. Alternately, we could apply a function $f$ to $\hat{\boldsymbol{n}}_1 \cdot \hat{\boldsymbol{n}}_2$ that is minimized at some $\hat{\boldsymbol{n}}_1 \cdot \hat{\boldsymbol{n}}_2 = \cos\theta_* \neq 1$. In this case it is important to retain the correct quadratic behavior independent of $\theta_*$. Such a function is not immediately obvious, and would not necessarily justify itself.

For these reasons, we opt against using the normals alone in constructing the bending stiffness.

### 4.4.3 Growth Rules

As mentioned above, growth steps for a hexamer-based model need not be as complicated as the growth steps for a trimer-based model. We need only the step of adding a new unit to an existing border edge. Because every interior vertex has exactly three units around it, we no longer need to delay the decision to join adjacent edges together. That is, if two units already meet at a vertex, and a new unit is added to one of the free edges, it can automatically be joined to both free edges emanating from that vertex. The only choices left to be made are (a) where to add the new unit, and (b) whether the unit should be a hexon or penton (or whether such transitions can be made after the fact by units with external edges, similar to the work by Zandi et al. [19]).

Recent work by Levandovsky and Zandi [1] has shown that allowing topologically distant segments of a growing capsid to reconnect is helpful in preventing the crevice failures we experienced in our triangular network growth model. Since we've removed the pentons from our model, we must instead set up the growth rules in such a way that the capsid will branch out and later rejoin topologically distant units when they come into close spatial proximity.

This branching requires small modifications to the interactions: specifically, we

must allow topologically distant vertices that are capable of joining to have a spatially short-range attraction, and border vertices (and possibly hexamer centers, as well) that cannot join to have a spatially short-range repulsion. With a smart choice of cut-off distance $x_0$ (probably the hexagon radius $r_0$) we should be able to use (4.3), along with an inverted version of the same potential for the repulsion, thus avoiding the introduction of additional parameters.

## 4.5 Fitting model parameters to experimental data

Our ultimate goal in developing a model is to grow an ensemble of capsids, as we did in Chapter 2. In particular, for the immature HIV model in §4.4, we would look at the positions, shapes and sizes of holes, in addition to the other measurements discussed in §2.3. While time did not permit carrying out such a growth simulation, we did carry out some preliminary work. Before a reasonable investigation of growth can begin, we need reasonable estimates of the numerical parameters. In this section we will outline several approaches to fitting these parameters to experimental data and end with preliminary results in §4.5.3.

### 4.5.1 Cryo-tomography data

We have received cryo-tomography data of immature HIV capsids from the Briggs lab [18]. Since the positions and angles between units are determined by the energy parameters $r_0$, $\ell$, and $\theta_*$, we should be able to find the parameters that are most likely to have produced the data (note that the one parameter with units of energy cannot be fit, except possibly with more sophisticated techniques comparing the measurement uncertainties with the cryogenic temperature). There are a number of ways we

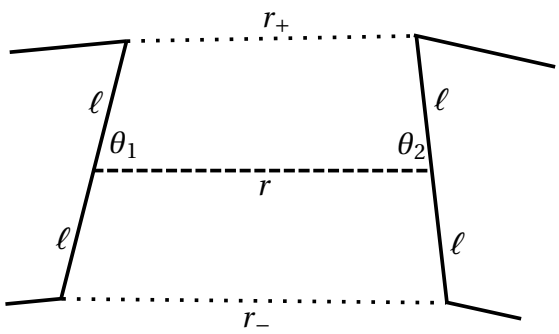



Figure 4.6: Side view of a pair of interacting edges. The distances between the centers of each edge is $r$, and the distances between the pairs of contacts are $r_+$ and $r_-$, which we square to get the energy.

might go about such a calculation.

In any approach, we must first reconstruct the topology. The data are specified in terms of the spatial position and orientation of each of several hundred hexameric units, but there is no connectivity information included. Moreover, the orientation is relative to a particular reference orientation (see §5.2). While we can identify the six-fold symmetry axis, there is still a free angle $\phi$, mapping $\frac{m\pi}{6} \to \frac{m\pi}{6} + \phi$ in (4.1), which corresponds to a rotation about the six-fold axis. We must fit this parameter as well, since the actual hexamers may not necessarily bind exactly at their corners.

### 4.5.2 Difficulties of fitting parameters

A naïve first attempt at this point is to fix the positions and orientations of all the hexamers to their measured values and minimize the energy as a function of the four parameters $r_0$, $\theta_*$, $\ell$, and $\phi$. The is rather unsatisfying, as it results always in $\ell \to 0$. This can be seen by considering a pair of interacting contact points on one pair of edges on a pair of interacting faces, as shown in Figure 4.6. Such a local picture eliminates the $r_0$ and $\theta_*$ parameters by specifying a single edge at a time. The energy due

to any such interacting pair is given by

$$E_{\text{pair}} = \frac{1}{2}k(r_1^2 + r_2^2) = k\Big(r^2 + \ell^2(\sin\theta_1 + \sin\theta_2)^2\Big), \tag{4.6}$$

which we can immediately see is minimized by $\ell = 0$. This is clearly the wrong approach, since we're minimizing the energy parameters while holding the positions fixed, while we should be doing the opposite.

The next attempt we might make while still fixing the positions is to compute the forces and torques and minimize the sums of squares. This requires an extra parameter in the form of a constant of proportionality $\rho$ between the forces and torques so that we can add them. Such a parameter is a length, and a natural choice would be something on the order of the size of the hexameric units.

Unfortunately, this plan fails in practice as well. We can show that the forces are independent of $\ell$, while the torques are a linear combination of $\ell$ and $r_0$ that vanishes at the energy minimum. It's not an unreasonable assumption that averaging over the near-minimum geometries will give an expected squared torque of

$$\langle \tau^2 \rangle \propto r_0^2 + \gamma \ell^2 \tag{4.7}$$

for some constant $\gamma$. If this is indeed the case, then we would expect that, regardless of the value of the proportionality constant $\rho$, the sum $\rho \sum \boldsymbol{F}^2 + \sum \boldsymbol{\tau}^2$ will again be minimized when $\ell = 0$.

### 4.5.3 Fit and results

Instead of minimizing energies or forces, we should instead consider what we are really trying to optimize, that is, the probability of the model given the data, $P(\text{M}|\text{D})$. While this is impossible to measure directly, we can instead assume that the model space is distributed with uniform probabilities and calculate $P(\text{D}|\text{M})$ by relaxing

the positions under any given set of energy parameters and computing the sum of squared distances between the original and relaxed positions and orientations[10].

Because each step in this optimization requires its own minimization, it is a rather slow process, but the condition number is reasonable, since we only have four degrees of freedom ($r_0$, $\theta_*$, $\ell$, and $\phi$). A simple downhill simplex algorithm is therefore sufficient to locate a reasonable minimum. The result of such an optimization is

$$r_0 \approx 5.6\text{nm} \tag{4.8}$$

$$\ell \approx 2\text{nm} \tag{4.9}$$

$$\theta_* \approx 5^\circ. \tag{4.10}$$

These numbers are relatively stable to changes in the precise positions of the reference points. Moreover, $\theta_*$ is the sloppiest, varying up to 30% depending on initial conditions, followed by $\ell$, which varied less, while $r_0$ was (as expected) completely stable.

Assuming a flat lattice, $r_0 = 5.6$nm gives an interhexamer spacing of 9.7nm, which is a bit larger than the accepted value of 8nm [18], but is exactly consistent with the average nearest-neighbor spacing in the provided structure.

We can get an idea of the angle $\theta_*$ be calculating that a spherical hexagon with circumradius $\theta_* = 5^\circ$ has a spherical area of 0.02, so that 630 units can fit on a full sphere, which is less than the $\approx 750$ units in each structure. Add to this the fact that the structures have holes and we see we can conclude that $\theta_*$ is likely overestimated.

The thickness $\ell = 2$nm translates to $\ell_\text{f}^2 = 3\text{nm}^2$ and (assuming a radius 50nm) a Foppl-von Kármán number $\gamma = (R/\ell_\text{f})^2 = 800$, putting well into the buckled regime.

[10] Again, a natural constant of proportionality arises by picking multiple reference points on each hexamer and computing just the distances between the original and relaxed reference points.

It is hard to gauge whether this is accurate, since we don't know the effect of holes on buckling. However, the structures don't *look* buckled, and it's conceivable that we underestimated $\ell$ because of incorrectly assigned neighbors or occasional outliers with large bend angles.

These results provide a starting point for a more detailed measurement of elastic parameters from cryo-electron tomographs, which can ultimately motivate a more accurate coarse-grained model of immature HIV assembly.

## BIBLIOGRAPHY


[1] A. Levandovsky and R. Zandi, "Nonequilibirum Assembly, Retroviruses, and Conical Structures", *Phys. Rev. Lett.* **102**, 198102 (2009).

[2] V. S. Reddy, H. A. Giesling, R. T. Morton, A. Kumar, C. B. Post, C. L. Brooks, and J. E. Johnson, "Energetics of Quasiequivalence: Computational Analysis of Protein-Protein Interactions in Icosahedral Viruses", *Biophys. J.* **74**, 546 (1998).

[3] R. Schwartz, P. W. Shor, P. E. Prevelige, and B. Berger, "Local Rules Simulation of the Kinetics of Virus Capsid Self-Assembly", *Biophys. J.* **75**, 2626 (1998).

[4] D. C. Rapaport, "Self-assembly of polyhedral shells: A molecular dynamics study", *Phys. Rev. E* **70**, 51905 (2004).

[5] M. F. Hagan and D. Chandler, "Dynamic Pathways for Viral Capsid Assembly", *Biophys. J.* **91**, 42 (2006).

[6] C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, "Transition path sampling and the calculation of rate constants", *J. Chem. Phys.* **108**, 1964 (1998).

[7] E. Weinan, W. Ren, and E. Vanden-Eijnden, "String method for the study of rare events", *Phys. Rev. B* **66**, 052301 (2002).

[8] J. A. Briggs, K. Grünewald, B. Glass, F. Förster, H. G. Kräusslich, and S. D. Fuller, "The Mechanism of HIV-1 Core Assembly: Insights from Three-Dimensional Reconstructions of Authentic Virions", *Structure* **14**, 15 (2006).

[9] J. Benjamin, B. K. Ganser-Pornillos, W. F. Tivol, W. I. Sundquist, and G. J. Jensen, "Three-dimensional structure of HIV-1 virus-like particles by electron cryotomography.", *J. Mol. Biol.* **346**, 577 (2005).

[10] J. A. G. Briggs, T. Wilk, R. Welker, H. G. Kräusslich, and S. D. Fuller, "Structural organization of authentic, mature HIV-1 virions and cores", *EMBO J.* **22**, 1707 (2003).

[11] E. R. Wright, J. B. Schooler, H. J. Ding, C. Kieffer, C. Fillmore, W. I. Sundquist, and G. J. Jensen, "Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells.", *EMBO J.* **26**, 2218 (2007).

[12] R. F. Bruinsma, W. M. Gelbart, D. Reguera, J. Rudnick, and R. Zandi, "Viral Self-Assembly as a Thermodynamic Process", *Phys. Rev. Lett.* **90**, 248101 (2003).

[13] H. D. Nguyen and C. L. B. III, "Generalized Structural Polymorphism in Self-Assembled Viral Particles", *Nano Lett.* **8**, 4574 (2008).

[14] D. L. D. Caspar and A. Klug, "Physical Principles In Construction Of Regular Viruses", *Cold Spring Harb. Symp. Quant. Biol.* **27**, 1 (1962).

[15] R. Twarock, "A tiling approach to virus capsid assembly explaining a structural puzzle in virology", *J. Theor. Biol.* **226**, 477 (2004).

[16] D. C. Rapaport, "Self-assembly of polyhedral shells: A molecular dynamics study", *Phys. Rev. E* **70**, 051905 (2004).

[17] P. Forrer, C. Chang, D. Ott, A. Wlodawer, and A. Pluckthun, "Kinetic Stability and Crystal Structure of the Viral Capsid Protein SHP", *J. Mol. Biol.* **344**, 179 (2004).

[18] J. A. G. Briggs, J. D. Riches, B. Glass, V. Bartonova, G. Zanetti, and H. G. Kräusslich, "Structure and assembly of immature HIV", *Proc. Nat. Acad. Sci. USA* **106**, 11090 (2009).

[19] R. Zandi, D. Reguera, R. F. Bruinsma, W. M. Gelbart, and J. Rudnick, "Origin of icosahedral symmetry in viruses", *Proc. Nat. Acad. Sci. USA* **101**, 15556 (2004).

CHAPTER 5

# GENERALIZED SPRINGS

We mentioned briefly in §4.1.1 the possibility of performing growth simulations with different potentials between the units. We now revisit this possibility in an effort to motivate our model parameters with experimental data, using atomistic simulations to fit the parameters of a very general quadratic potential. In particular, we favor increasing the number of fit parameters over injecting any preconceived notions of what this form should look like.

An unfortunate consequence of increasing the number of parameters is that we need a considerable amount of formalism to deal with them reasonably. Thus, we devote this chapter to establishing the generalized spring formalism, while actually applying it in Chapters 6 and 7 (although these chapters are fully comprehensible without the details of this chapter).

We begin in §5.1 by defining our general harmonic potential in terms of the six-dimensional *misconfiguration* vector, which we proceed to define in §5.2, including the various gauge choices (this is all that was directly used in Chapter 6). In §5.3, we write these gauge transformations explicitly. Finally, we conclude in §5.4 with a derivation of analytical spring composition laws, giving the total stiffness for a pair of generalized springs connected in series.

## 5.1 Preliminary formalism

Consider the most general harmonic interaction between two rigid bodies in three-space. Each rigid body has six degrees of freedom: three translational and three orientational, giving a total of twelve degrees of freedom for the composite system. Of these twelve, six are global translations and rotations, leaving six internal (relative)

degrees of freedom to describe the configuration of the spring, $\mathbb{r}_i$, $i = 1 \ldots 6$ (we'll use blackboard bold to represent 6-vectors). If we assume a minimum-energy configuration at $\overline{\mathbb{r}}$ then the simplest harmonic potential is simply

$$\mathscr{F} = \frac{1}{2}(\mathbb{r} - \overline{\mathbb{r}})^{\mathsf{T}} \mathbf{K} (\mathbb{r} - \overline{\mathbb{r}}), \tag{5.1}$$

where $\mathbf{K}$ is a $6 \times 6$ generalized stiffness tensor.

So how do we compute the *misconfiguration* $\delta\mathbb{r} \equiv$ "$\mathbb{r} - \overline{\mathbb{r}}$"? This is unfortunately complicated primarily by the presence of rotations, both because they are nonlinear (so we can't just subtract) and because we need the misconfiguration to be invariant under global rotations of the system.

In this section, therefore, we will go on to provide a precise definition for $\delta\mathbb{r}$ for small deviations from equilibrium. A primary difficulty in this discussion is notation. For the sake of clarity, we will be very pedantic, but the short story can be read off of the equations in Table **??**.

### 5.1.1 Rotation representations

The first issue at hand is our representation of rotations. There are a number of different ways we can look at rotations. Most generally, a rotation $\boldsymbol{\Omega}$ is any inner product-preserving linear automorphism over vectors $\boldsymbol{r} \in \mathbb{R}^3$. That is, $\boldsymbol{\Omega}(\boldsymbol{u} + \boldsymbol{v}) = \boldsymbol{\Omega}\boldsymbol{u} + \boldsymbol{\Omega}\boldsymbol{v}$ and $\boldsymbol{\Omega}\boldsymbol{u} \cdot \boldsymbol{\Omega}\boldsymbol{v} = \boldsymbol{u} \cdot \boldsymbol{v}$. The most familiar representation is SO(3), the set of orthogonal $3 \times 3$ matrices.

We can describe a rotation by writing an axis $\hat{\boldsymbol{\theta}}$ and an amount of rotation $|\boldsymbol{\theta}|$ about that axis: $\boldsymbol{\Omega} = \mathsf{R}(\boldsymbol{\theta})$, or inversely, $\boldsymbol{\theta} = \boldsymbol{R}^{-1}(\boldsymbol{\Omega})$. We can translate between the matrix and axis representations as follows:

$$\mathsf{R}(\boldsymbol{\theta}) = \exp[\boldsymbol{\theta}\times] = \hat{\boldsymbol{\theta}} \otimes \hat{\boldsymbol{\theta}} + \cos|\boldsymbol{\theta}| \left(\mathbf{1} - \hat{\boldsymbol{\theta}} \otimes \hat{\boldsymbol{\theta}}\right) + \sin|\boldsymbol{\theta}| \, [\hat{\boldsymbol{\theta}}\times], \tag{5.2}$$

Table 5.1: Rigid body notation summary. All quantities below the gauge choice $\mathbf{G}$ are (implicitly) dependent on that choice. We drop the subscripts in the bottom section since they obscure the expressions without adding any value. $\lambda$ in the definition of $\boldsymbol{\delta\omega}$ is 0, 1, or $\frac{1}{2}$, depending on the gauge, as described in §5.2.4. The simplest case, the $\alpha$-gauge, is given by $\mathbf{G} = \boldsymbol{\Theta}_\alpha$ and $\lambda = 0$.

| Notation | Name | Type | Definition |
|---|---|---|---|
| $\alpha, \beta$ | Body/frame index | | 0, 1, 2, … |
| $\mu$ | Spring/pair index | | A, B, … |
| $\mathsf{R}(\boldsymbol{\theta})$ | Rotation | SO(3) | See (5.2) |
| $\boldsymbol{r}_i^{(\alpha)}$ | Constituent points | $\mathbb{R}^3$ | |
| $\boldsymbol{\Omega}_{\alpha\beta}$ | Transformation | SO(3) | $\boldsymbol{r}_i^{(\alpha)} \equiv \boldsymbol{\Omega}_{\alpha\beta}\boldsymbol{r}^{(\beta)}$ |
| $\boldsymbol{R}_\alpha$ | Absolute translation | $\mathbb{R}^3$ | $\langle \boldsymbol{r}_i^{(\alpha)} \rangle$ |
| $\boldsymbol{\Theta}_\alpha$ | Absolute orientation | SO(3) | $\boldsymbol{\Omega}_{\alpha 0}$ |
| $\mathbb{R}_\alpha$ | Absolute configuration | $\mathbb{R}^3 \times$ SO(3) | $(\boldsymbol{R}_\alpha, \boldsymbol{\Theta}_\alpha)$ |
| $\mathbf{G}$ | Gauge choice | SO(3) | See (5.10) |
| $\boldsymbol{r}_{\alpha\beta}$ | Relative position | $\mathbb{R}^3$ | $\mathbf{G}(\boldsymbol{R}_\beta - \boldsymbol{R}_\alpha)$ |
| $\boldsymbol{\omega}_{\alpha\beta}$ | Relative orientation | SO(3) | $\mathbf{G}\boldsymbol{\Theta}_\alpha^\mathsf{T}\boldsymbol{\Theta}_\beta\mathbf{G}^\mathsf{T}$ |
| $\mathbb{r}_{\alpha\beta}$ | Relative configuration | $\mathbb{R}^3 \times$ SO(3) | $(\boldsymbol{r}_{\alpha\beta}, \boldsymbol{\omega}_{\alpha\beta})$ |
| $\overline{\boldsymbol{r}}_{\alpha\beta}$ | Preferred position | $\mathbb{R}^3$ | $\langle \boldsymbol{r}_{\alpha\beta} \rangle$ |
| $\overline{\boldsymbol{\omega}}_{\alpha\beta}$ | Preferred orientation | SO(3) | $\langle \boldsymbol{\omega}_{\alpha\beta} \rangle$ |
| $\overline{\mathbb{r}}_{\alpha\beta}$ | Preferred configuration | $\mathbb{R}^3 \times$ SO(3) | $(\overline{\boldsymbol{r}}_{\alpha\beta}, \overline{\boldsymbol{\omega}}_{\alpha\beta})$ |
| $\boldsymbol{\delta r}$ | Mistranslation | $\mathbb{R}^3$ | $\boldsymbol{r} - \overline{\boldsymbol{r}}$ |
| $\boldsymbol{\delta\omega}$ | Misorientation | SO(3) | $\overline{\boldsymbol{\omega}}^{-1+\lambda}\boldsymbol{\omega}\overline{\boldsymbol{\omega}}^{-\lambda}$ |
| $\boldsymbol{\delta\vartheta}$ | | $\mathbb{R}^3$ | $\boldsymbol{R}^{-1}(\boldsymbol{\delta\omega})$ |
| $\mathbb{\delta r}$ | Misconfiguration | $\mathbb{R}^6$ | $(\boldsymbol{\delta r}, \boldsymbol{\delta\vartheta})$ |

where $[\boldsymbol{\theta}\times]$ is the skew-symmetric form of $\boldsymbol{\theta}$, that is, a $3\times 3$ matrix that multiplies $[\boldsymbol{\theta}\times]\boldsymbol{r} = \boldsymbol{\theta}\times\boldsymbol{r}$. Note also that $\mathrm{Tr}\,\mathsf{R}(\boldsymbol{\theta}) = 1 + 2\cos|\boldsymbol{\theta}|$.

A third representation, particularly convenient for computers, is the unit quaternions, which are isomorphic to SU(2). Here we represent a rotation as $\mathsf{q} = \rho\mathbf{1} - \mathsf{i}\boldsymbol{\zeta}\cdot\sigma$, where $\mathsf{i}$ is the imaginary unit, $\sigma$ is the vector of Pauli matrices, and $\rho^2 + |\boldsymbol{\zeta}^2| = 1$. To construct a quaternion from a rotation through the axis $\boldsymbol{\theta}$, we write $\rho = \cos\frac{|\boldsymbol{\theta}|}{2}$ and $\boldsymbol{\zeta} = \hat{\boldsymbol{\theta}}\sin\frac{|\boldsymbol{\theta}|}{2}$. (The unit quaternions in fact are a double cover of the rotations, since $|\boldsymbol{\theta}| \to |\boldsymbol{\theta}| + 2\pi$ gives $\mathsf{q} \to -\mathsf{q}$, to represent the same rotation.) To apply a rotation to a vector $\boldsymbol{v}$ we (matrix) multiply $\mathsf{i}\sigma\cdot\boldsymbol{v}' = \mathsf{q}(\mathsf{i}\sigma\cdot\boldsymbol{v})\mathsf{q}^\dagger$. Composition is achieved by quaternion multiplication:

$$(\rho\mathbf{1} - \mathsf{i}\boldsymbol{\zeta}\cdot\sigma)(\rho'\mathbf{1} - \mathsf{i}\boldsymbol{\zeta}'\cdot\sigma) = (\rho\rho' - \boldsymbol{\zeta}\cdot\boldsymbol{\zeta}') - \mathsf{i}(\rho\boldsymbol{\zeta}' + \rho'\boldsymbol{\zeta} + \boldsymbol{\zeta}\times\boldsymbol{\zeta}'). \tag{5.3}$$

Alternately, quaternions can be viewed as adjoining a second imaginary unit $\mathsf{j}$, with $\mathsf{j}^2 = -1$, and defining $\mathsf{k} = \mathsf{ij} = -\mathsf{ji}$, such that $\mathsf{q} = \rho + \mathsf{i}\zeta_x + \mathsf{j}\zeta_y + \mathsf{k}\zeta_z$.

At small angles $\boldsymbol{\theta} \to \mathbf{0}$, we can make several simplifying approximations. In particular, the matrix representation linearizes to

$$\mathsf{R}(\boldsymbol{\theta} \to \mathbf{0}) = \mathbf{1} + [\boldsymbol{\theta}\times]. \tag{5.4}$$

Thus, to linear order in this limit, composition is commutative vector addition, $\mathsf{R}(\boldsymbol{\theta})\mathsf{R}(\boldsymbol{\theta}') = \mathsf{R}(\boldsymbol{\theta} + \boldsymbol{\theta}')$. The quaternion representation reduces to $\rho = 1$, $\boldsymbol{\zeta} = \frac{1}{2}\boldsymbol{\theta}$.

### 5.1.2 Passive and active transformations

One particularly confusing aspect of transformations, and specifically rotations, is the difference between passive and active transformations. We will use *active transformations* everywhere, for reasons motivated below.

Active transformations are transformations on objects. That is, suppose we have an object consisting of points $\{\boldsymbol{r}_i^{(\alpha)}\}$ and a reference copy of the same object with points $\{\boldsymbol{r}_i^{(0)}\}$. An active transformation from the reference object to the actual object is a transformation $\mathsf{A}$ such that $\boldsymbol{r}_i^{(\alpha)} = \mathsf{A}\boldsymbol{r}_i^{(0)}$.

Passive transformations, on the other hand, are transformations of reference frames. Given the actual and reference copies above, a passive transformation $\mathsf{P}$ from the reference frame to the actual frame must satisfy $\boldsymbol{x}^{(\alpha)} = \mathsf{P}\boldsymbol{x}^{(0)}$, where we define $\boldsymbol{x}^{(\alpha)} \cdot \boldsymbol{r}_i^{(\alpha)} = \boldsymbol{x}^{(0)} \cdot \boldsymbol{r}_i^{(0)}$ for all $i$ (disregarding any translations[1]).

The upshot is that these transformations are inverses: $\mathsf{P} = \mathsf{A}^{-1}$. Or alternately, transforming from the reference object to the actual object is equivalent to transforming from the actual frame to the reference frame.

There are advantages to each representation. Active transformations are more concrete to reason about, since we think about relationships between objects daily, while we think about relationships between differing points of view only when we have no other recourse. Another advantage to active transformations is that we need only ever consider a single reference frame, given by the global coordinate system. This obviates the difficulties associated with bookkeeping which frame a given vector belongs to, while a passive approach necessitates keeping track of which object/frame a vector belongs to.

The advantage of passive transformations is slightly more subtle. Because we will be dealing with pairs of objects in which the relationship between the pair is of complete importance, and any global transformations are irrelevant, it is helpful to think

[1] We could be particularly thorough and introduce a new notation $[\boldsymbol{r}]_\alpha$ to represent the coordinates of the vector $\boldsymbol{r}$ when represented in the reference frame $\alpha$. But in the absence of translations, the components of $[\boldsymbol{r}]_\alpha$ are simply inner products with the coordinate basis vectors.

about expressing quantities in each object's reference frame: it's much more expressive to say that "B is in front of A and A is in front of B", rather than "B is east of A, A is facing east, and B is facing west".

References to "front" and "facing", of course, imply that we have figured out a way to glue a set of coordinate axes onto our objects, and this is a gauge freedom in both representations.

Thus, in the interest of overall comprehensibility, we will assume an active transformation picture from this point onward[2].

## 5.2 Absolute and relative configurations

As we can see from Table **??**, there is a lot of ground to cover here. We start by defining the absolute configuration $\mathbb{R}_\alpha$ of a single body given a reference configuration (§5.2.1), along with the gauge freedoms of rotating either the reference or the object (§5.2.2). We then go on to define the relative configuration $\mathbb{r}_{\alpha\beta}$ between two objects (§5.2.3) and the relevant gauge choices (§5.2.4). Finally, we assume a preferred (relative) configuration $\overline{\mathbb{r}}_{\alpha\beta}$ and discuss how to express the misconfiguration $\delta\mathbb{r}_{\alpha\beta}$ in any particular gauge (§5.2.5). As an appendix, we consider a more natural gauge with one translational and five rotational degrees of freedom (§5.2.6).

### 5.2.1 One body: absolute configuration

Consider first a single object $\alpha$ in space, composed of points $\{\boldsymbol{r}_i^{(\alpha)}\}$. Its position, or *absolute translation*, is simply the center of mass $\boldsymbol{R}_\alpha = \frac{1}{N}\sum \boldsymbol{r}_i^{(\alpha)}$. To define its *absolute orientation* $\boldsymbol{\Theta}_\alpha$, we need a reference to compare to. Take $\{\boldsymbol{r}_i^{(0)}\}$ to be a reference copy

[2]Our original computations were all done with passive transformations, but since the results are equivalent up to a matrix transpose, they are still valid.

of the object (with center of mass $\boldsymbol{R}^{(0)}$ at the origin). Then we can determine the rotation matrix $\boldsymbol{\Omega}_{\alpha 0}$ such that

$$\boldsymbol{r}_i^{(\alpha)} = \boldsymbol{R}_\alpha + \boldsymbol{\Omega}_{\alpha 0} \boldsymbol{r}_i^{(0)} \tag{5.5}$$

for all $i$. It's possible (quite likely, in fact) that there is no rotation $\boldsymbol{\Omega}_{\alpha 0}$ that satisfies (5.5) for all $i$. In this case, we can define a

$$\chi^2 = \sum_i \left( \boldsymbol{R}_\alpha + \boldsymbol{\Omega}_{\alpha 0} \boldsymbol{r}_i^{(0)} - \boldsymbol{r}_i^{(\alpha)} \right)^2 \tag{5.6}$$

to minimize over $\boldsymbol{\Omega}_{\alpha 0}$.

While we can define such rotation matrices between any pair of objects, the case of a rotation to an actual object from its *reference* copy is the special case we define as absolute orientation, and we thus write

$$\boldsymbol{\Theta}_\alpha = \boldsymbol{\Omega}_{\alpha 0}. \tag{5.7}$$

Taken together, the absolute translation and orientation comprise what we will call the *absolute configuration*:

$$\mathbb{R}_\alpha = (\boldsymbol{R}_\alpha, \boldsymbol{\Theta}_\alpha). \tag{5.8}$$

### 5.2.2 Gauge freedoms: reference and scene rotations

Even with a single body there are a number of gauge freedoms we can begin to explore. Suppose we take a rotated reference configuration $0'$, with $\boldsymbol{r}_{0'} = \boldsymbol{\Omega}_{0'0} \boldsymbol{r}_0$. (Note we'll drop the explicit component point index subscript $i$ in favor of using the subscripts to represent objects, while still assuming the equations hold for all component points). The result of this *reference rotation* is that

$$\boldsymbol{\Theta}_\alpha \to \boldsymbol{\Theta}'_\alpha = \boldsymbol{\Omega}_{\alpha 0'} = \boldsymbol{\Omega}_{\alpha 0} \boldsymbol{\Omega}_{00'} = \boldsymbol{\Theta}_\alpha \boldsymbol{\Omega}_{0'0}^{\mathsf{T}}. \tag{5.9}$$

Alternately, we can rotate the object by a matrix $\mathbf{G} = \mathbf{\Omega}_{\alpha'\alpha}$. The result of this *scene rotation* is

$$\mathbf{\Theta}_\alpha \to \mathbf{\Theta}'_\alpha = \mathbf{\Omega}_{\alpha' 0} = \mathbf{\Omega}_{\alpha'\alpha}\mathbf{\Omega}_{\alpha 0} = \mathbf{G}\mathbf{\Theta}_\alpha. \tag{5.10}$$

Obviously displacement vectors are also affected by a scene rotation:

$$\boldsymbol{r} \to \boldsymbol{r}' = \mathbf{G}\boldsymbol{r}. \tag{5.11}$$

This matrix $\mathbf{G}$ will provide the main important gauge freedom.[3]

### 5.2.3 Two bodies: relative configuration

Consider next two bodies, 1 and 2. This adds a significant amount of complexity, since we can now look at the *relative* degrees of freedom between the two objects. The *relative displacement* is straightforward in the global frame:

$$\boldsymbol{r}_{12} = \boldsymbol{R}_2 - \boldsymbol{R}_1. \tag{5.12}$$

To reason about the *relative orientation*, we will first assume that the two objects are identical and share a common reference, so that we have only three objects to consider. We can then define the relative orientation $\boldsymbol{\omega}_{12}$ to be what the absolute orientation of 2 would be *if 1 were the reference body*. That is[4],

$$\boldsymbol{\omega}_{12} = \mathbf{\Omega}_{21} = \mathbf{\Omega}_{20}\mathbf{\Omega}_{01} = \mathbf{\Theta}_2\mathbf{\Theta}_1^{\mathsf{T}}. \tag{5.13}$$

In the global frame, $\boldsymbol{r}_{12}$ and $\boldsymbol{\omega}_{12}$ are not particularly useful since they are not invariant

[3] It should be clear at this point that a rotation of the global reference frame (say a rotation $\mathbf{G}'$ to both reference and actual copies of the object) would have a net effect of conjugating $\mathbf{\Theta}_\alpha \to \mathbf{G}'\mathbf{\Theta}_\alpha\mathbf{G}'^{\mathsf{T}}$.

[4] In general $\mathbf{\Omega}_{\alpha\gamma} = \mathbf{\Omega}_{\alpha\beta}\mathbf{\Omega}_{\beta\gamma}$, since $\mathbf{\Omega}_{\beta\gamma}$ first rotates $\gamma$ to $\beta$, and then $\mathbf{\Omega}_{\alpha\beta}$ rotates $\beta$ to $\alpha$.

under a global scene rotation $\mathbf{G}$. Recall from (5.10) that a scene rotation transforms

$$(\boldsymbol{\Theta}_1, \boldsymbol{\Theta}_2) \longrightarrow (\mathbf{G}\boldsymbol{\Theta}_1, \mathbf{G}\boldsymbol{\Theta}_2) \tag{5.14}$$

$$(\boldsymbol{R}_1, \boldsymbol{R}_2) \longrightarrow (\mathbf{G}\boldsymbol{R}_1, \mathbf{G}\boldsymbol{R}_2)\,, \tag{5.15}$$

and thus

$$\boldsymbol{r}_{12} = \mathbf{G}(\boldsymbol{R}_2 - \boldsymbol{R}_1) \tag{5.16}$$

$$\boldsymbol{\omega}_{12} = \mathbf{G}\boldsymbol{\Theta}_2\boldsymbol{\Theta}_1^{\mathsf{T}}\mathbf{G}^{\mathsf{T}}. \tag{5.17}$$

We then define the *relative configuration* as the pair

$$\mathbb{r}_{12} = (\boldsymbol{r}_{12}, \boldsymbol{\omega}_{12})\,. \tag{5.18}$$

### 5.2.4 Gauge choice: body-centric and symmetric gauges

The choice of $\mathbf{G}$ is effectively a choice of which reference frame to measure the configuration (translation and orientation) relative to. The choice $\mathbf{G} = \mathbf{1}$ corresponds to measuring in the global reference frame, which is unnatural because it is effectively an arbitrary choice. A much more natural choice is the 1-gauge, $\mathbf{G} = \boldsymbol{\Theta}_1^{\mathsf{T}}$, which leaves

$$\boldsymbol{\omega}_{12}^{(1)} = \boldsymbol{\Theta}_1^{\mathsf{T}}\boldsymbol{\Theta}_2, \tag{5.19}$$

and

$$\boldsymbol{r}_{12}^{(1)} = \boldsymbol{\Theta}_1^{\mathsf{T}}\boldsymbol{r}_{12}. \tag{5.20}$$

Note that actual global rotations no longer have any effect, since any change in $\boldsymbol{\Theta}_1$ is reflected by a change in our $\mathbf{G}$, so that (5.19) as written is independent of any further transformations $\boldsymbol{\Theta}_\alpha \to \mathbf{G}'\boldsymbol{\Theta}_\alpha$ because the $\mathbf{G}'^{\mathsf{T}}$ and $\mathbf{G}'$ will cancel.

In addition, we can look at the 2-gauge, $\mathbf{G} = \boldsymbol{\Theta}_2^{\mathsf{T}}$. While the relative displacement vector is of course rotated into 2's reference frame, $\boldsymbol{r}_{12}^{(2)} = \boldsymbol{\Theta}_2^{\mathsf{T}}\boldsymbol{r}_{12}$, we find somewhat

surprisingly that $\boldsymbol{\omega}_{12}^{(2)} = \boldsymbol{\omega}_{12}^{(1)}$ is unchanged. We can reason that, because the matrix norm $\|\boldsymbol{\omega}_{12}\|$ is an invariant under conjugation, $\boldsymbol{\omega}_{12}$ occupies a lower-dimensional space than $\mathbf{G}$. In fact, for any $\boldsymbol{\omega}_{12}$, the subgroup of rotations coaxial with $\boldsymbol{\omega}_{12}$ leaves the relative orientation invariant.

This suggests there might be another, more appropriate gauge in between the two just discussed, that also leaves $\boldsymbol{\omega}_{12}$ invariant. Consider $\mathbf{G} = \boldsymbol{\omega}_{12}^{-1/2}\boldsymbol{\Theta}_1^{\mathsf{T}}$ (taking $\boldsymbol{\omega}_{12} = \boldsymbol{\Theta}_1^{\mathsf{T}}\boldsymbol{\Theta}_2$). This transformation results in absolute orientations

$$(\boldsymbol{\Theta}_1, \boldsymbol{\Theta}_2) \longrightarrow \left(\boldsymbol{\omega}_{12}^{-1/2}, \boldsymbol{\omega}_{12}^{1/2}\right), \tag{5.21}$$

which we will call the *symmetric gauge*.

### 5.2.5 Preferred configuration and misconfiguration

Now that we have defined relative configuration, we can assume a particular relative configuration $\overline{\mathbb{r}}_{12} = \left(\overline{\boldsymbol{r}}_{12}, \overline{\boldsymbol{\omega}}_{12}\right)$ is preferred for a given pair of objects[5]. Then for the configuration $\mathbb{r}_{12}$ at some snapshot we can define the *misconfiguration*

$$\delta\mathbb{r}_{12} = (\boldsymbol{\delta r}_{12}, \boldsymbol{\delta\omega}_{12})\,. \tag{5.22}$$

The *mistranslation* $\boldsymbol{\delta r}_{12} = \boldsymbol{r}_{12} - \overline{\boldsymbol{r}}_{12}$ is straightforward. The *misorientation* $\boldsymbol{\delta\omega}_{12}$, on the other hand, is not so simple, since rotations don't commute (and $\overline{\boldsymbol{\omega}}_{12}$ is not necessarily small). Moreover, the correct choice for $\boldsymbol{\delta\omega}_{12}$ in fact depends on our gauge choice.

[5]Computing the average $\langle\boldsymbol{\omega}\rangle$ of a set of rotation matrices is non-trivial. A simple component-wise matrix average is almost always non orthogonal. We can find the nearest orthogonal matrix to the average by solving the orthogonal Procrustes problem: given a matrix $\mathbf{M}$ we write its singular value decomposition $\mathbf{M} = \mathbf{U}\Sigma\mathbf{V}^{\dagger}$ where $\mathbf{U}$ and $\mathbf{V}$ are unitary and $\Sigma$ is diagonal. Then $\mathbf{R} = \mathbf{U}\mathbf{V}^{\dagger}$ is the orthogonal matrix such that the Frobenius norm $\|\mathbf{R} - \mathbf{M}\|_{\mathrm{F}}$ is minimized.

Consider first the 1-gauge. This basically amounts to an assumption that body 1 is fixed, so any deviation from the preferred configuration can be interpreted as a movement of body 2. Suppose then that $\boldsymbol{\Theta}_1$ is given, so that the preferred orientation of body 2 is $\boldsymbol{\Theta}_{\overline{2}} = \boldsymbol{\Theta}_1\overline{\boldsymbol{\omega}}_{12}$. Now if the actual orientation of body 2 is $\boldsymbol{\Theta}_2$ then the misorientation is naturally $\boldsymbol{\Omega}_{\overline{2}2}$. In our 1-gauge, with $\mathbf{G} = \boldsymbol{\Theta}_1^{\mathsf{T}}$, we find

$$\boldsymbol{\delta\omega}_{12}^{(1)} = \boldsymbol{\omega}_{\overline{2}2} = \boldsymbol{\Theta}_1^{\mathsf{T}}\boldsymbol{\Theta}_2\boldsymbol{\Theta}_{\overline{2}}^{\mathsf{T}}\boldsymbol{\Theta}_1 = \boldsymbol{\omega}_{12}\overline{\boldsymbol{\omega}}_{12}^{\mathsf{T}}. \tag{5.23}$$

On the other hand, in the 2-gauge, we are assuming that body 2 is fixed, so that the misorientation is given by

$$\boldsymbol{\delta\omega}_{12}^{(2)} = \boldsymbol{\omega}_{1\overline{1}} = \boldsymbol{\Theta}_2^{\mathsf{T}}\boldsymbol{\Theta}_1\boldsymbol{\Theta}_{\overline{1}}^{\mathsf{T}}\boldsymbol{\Theta}_2 = \overline{\boldsymbol{\omega}}_{12}^{\mathsf{T}}\boldsymbol{\omega}_{12}, \tag{5.24}$$

just the opposite of (5.23).

The symmetric gauge is slightly more complicated since we no longer assume either body is fixed. On the other hand, we do take the deviation of each body to be equal and opposite. Thus, if the "preferred" orientations are

$$\left(\mathbf{G}\boldsymbol{\Theta}_{\overline{1}}, \mathbf{G}\boldsymbol{\Theta}_{\overline{2}}\right) = \left(\overline{\boldsymbol{\omega}}_{12}^{-1/2}, \overline{\boldsymbol{\omega}}_{12}^{1/2}\right) \tag{5.25}$$

then we have actual orientations

$$(\mathbf{G}\boldsymbol{\Theta}_1, \mathbf{G}\boldsymbol{\Theta}_2) = \left(\boldsymbol{\omega}_{12}^{-1/2}, \boldsymbol{\omega}_{12}^{1/2}\right) = \left(\boldsymbol{\delta\omega}_{12}^{-1/2}\overline{\boldsymbol{\omega}}_{12}^{-1/2}, \boldsymbol{\delta\omega}_{12}^{1/2}\overline{\boldsymbol{\omega}}_{12}^{1/2}\right), \tag{5.26}$$

giving a misorientation

$$\boldsymbol{\delta\omega}_{12}^{(\mathrm{s})} = \overline{\boldsymbol{\omega}}_{12}^{-1/2}\boldsymbol{\omega}_{12}\overline{\boldsymbol{\omega}}_{12}^{-1/2}, \tag{5.27}$$

which is significantly less convenient to deal with.

On the other hand, if the actual configuration is sufficiently close to the preferred configuration, then the misconfiguration will be small and the gauge choice is irrelevant since we can write the misorientation as a small vector $\boldsymbol{\delta\vartheta}$, with $\boldsymbol{\delta\omega}_{12} = \mathsf{R}(\boldsymbol{\delta\vartheta}_{12})$,

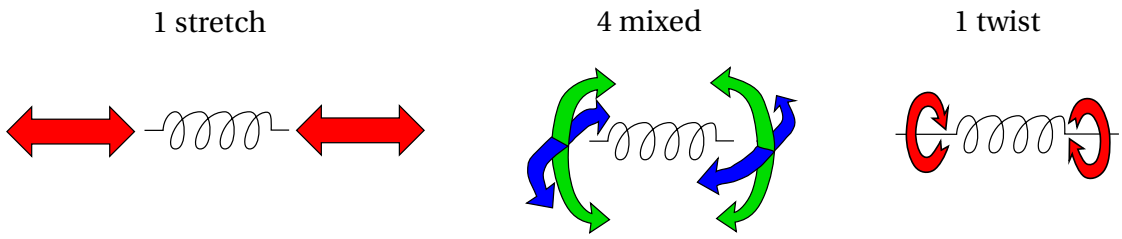



Figure 5.1: The six relative degrees of freedom between a pair of rigid bodies can be decomposed into one pure translation, called *stretch*, one pure rotation about the stretch axis, called *twist*, and four mixed degrees of freedom. The mixed degrees can be thought of as *shears* or *folds*, depending on whether the bodies move symmetrically or antisymmetrically.

and the misconfiguration as a 6-vector $\delta\mathfrak{r}_{12} = (\boldsymbol{\delta r}_{12}, \boldsymbol{\delta\theta}_{12})$. This is the simplest representation of the misconfiguration, which we'll call the 3+3 representation since it involves three translational and three orientational degrees of freedom.

### 5.2.6 The 1+5 representation

So far we have simply assumed that, because each absolute configuration had three translational and three orientational degrees of freedom, and the global translations and rotations had three of each as well, that the relative configuration would necessarily be left with three translational and three orientational degrees of freedom as well. This assumption led us to the 3+3 representation, described in full detail above.

We found, however that while three-dimensional translations made the math more straightforward, they weren't as natural as we might have liked. In particular, we ended up with a strange gauge choices where rotations might look like translations. If we instead think about the principal relative motions in everyday language, we come up with three different types of motions, as shown in Figure 5.1: a pure translation, a pure rotation, and four mixed degrees of freedom. The 3+3 representation decomposed these mixed degrees of freedom into two translations and two rotations, but

we might do better to leave them all as rotations, leading to the 1+5 picture.

In the 1+5 picture, only stretches—translations parallel to the interbody axis ($\overline{\boldsymbol{r}}_{12}$)—are written as translations, with units of distance, while the other two translations ("shears") are written as rotations in analogue with the "fold" rotations.[6]

To transform our 3+3 picture into a 1+5 picture, we start by finding a rotation $\mathbf{G}$ and a vector $\boldsymbol{C}$ such that

$$\mathbf{G}(\boldsymbol{R}_1 - \boldsymbol{C}) = \frac{d}{2}\hat{\boldsymbol{z}}, \tag{5.28}$$

$$\mathbf{G}(\boldsymbol{R}_2 - \boldsymbol{C}) = -\frac{d}{2}\hat{\boldsymbol{z}}. \tag{5.29}$$

That is, we line up the two objects along the $z$-axis. There is a one-dimensional family of rotations that will satisfy this condition: rotations about the $z$-axis will leave the condition invariant, so we have a gauge freedom in how we choose this rotation component (we might look for optimal alignment of one body with its reference orientation, or a mid-way point between two such alignments).

Once both the reference configurations and the instantaneous snapshots are transformed to satisfy our $z$-axis constraint, we can measure three quantities: first, the (stretch) distance $d$ between the bodies; second and third, the orientations $\boldsymbol{\Theta}_1$ and $\boldsymbol{\Theta}_2$ that best align each body with its reference. If we write these rotation matrices as rotation axes $\boldsymbol{\theta}_\alpha$ where $\boldsymbol{\Theta}_\alpha = \mathsf{R}(\boldsymbol{\theta}_\alpha)$, then we can decompose them into $\theta_{\alpha x}$, $\theta_{\alpha y}$, and $\theta_{\alpha z}$. Next, we construct the symmetric and antisymmetric combinations of

[6] We could potentially also talk about a 5+1 picture, but this would have all the disadvantages of the 1+5 picture with none of the benefits.

these components:

$$\psi_x = \frac{1}{2}\left(\theta_{1x} + \theta_{2x}\right), \tag{5.30}$$

$$\psi_y = \frac{1}{2}\left(\theta_{1y} + \theta_{2y}\right), \tag{5.31}$$

$$\theta_z = \frac{1}{2}\left(\theta_{1z} + \theta_{2z}\right), \tag{5.32}$$

$$\phi_x = \theta_{2x} - \theta_{1x}, \tag{5.33}$$

$$\phi_y = \theta_{2y} - \theta_{1y}, \tag{5.34}$$

$$\tau_z = \theta_{2z} - \theta_{1z}. \tag{5.35}$$

The symmetric $x$ and $y$ combinations, $\psi_{x,y}$ are the shears, while the corresponding antisymmetric combinations, $\phi_{x,y}$ are the folds. The antisymmetric $z$ combination $\tau_z$ is the twist, while $\theta_z$ represents our gauge freedom. It becomes apparent that transforming $\theta_z \to \theta_z + \theta$ has the effect of rotating $\psi_x \to \psi_x \cos\theta - \psi_y \sin\theta$ and similar for $\psi_y$ and $\phi_{x,y}$, while leaving $\tau_z$ unchanged. We can make this gauge choice of $\theta$ irrelevant by changing the model to impose a cylindrical symmetry about the interbody axis (reducing our degree of freedom space from six- to four-dimensional).

## 5.3 Tensor transformations

Here we define transformation matrices for each of the gauge transformations we introduced in §5.2, as well as any simultaneous modifications we need to make to $\overline{\boldsymbol{r}}_{12}$ and $\overline{\boldsymbol{\omega}}_{12}$ when transforming between gauges.

### 5.3.1 Covariance

For any given gauge transformation $\delta\mathfrak{r} \to \delta\mathfrak{r}'$, we would like to know how to transform the stiffness (**K**) and diffusion (**D**) tensors. We will define these tensors more

completely in Chapter 6. For now, we will start from (5.1), taking the misconfiguration $\boldsymbol{x}-\overline{\boldsymbol{x}} \to \delta\mathfrak{r}$ to be a 6-component column vector. Since our transformation is linear at small $\delta\mathfrak{r}$, we can write

$$\delta\mathfrak{r}' = \Lambda\delta\mathfrak{r} \tag{5.36}$$

for a $6\times 6$ matrix $\Lambda$ with units dictated by the particular gauge transformation. Our free energy (5.1) becomes

$$\mathscr{T} = \frac{1}{2}\delta\mathfrak{r}^{\mathsf{T}}\mathbf{K}\delta\mathfrak{r} = \frac{1}{2}\left(\delta\mathfrak{r}^{\mathsf{T}}\Lambda^{\mathsf{T}}\right)\left(\Lambda^{-\mathsf{T}}\mathbf{K}\Lambda^{-1}\right)(\Lambda\delta\mathfrak{r}) = \frac{1}{2}\delta\mathfrak{r}'^{\mathsf{T}}\mathbf{K}'\delta\mathfrak{r}', \tag{5.37}$$

so we see that $\mathbf{K}$ is covariant with $\delta\mathfrak{r}$:

$$\mathbf{K}' = \Lambda^{-\mathsf{T}}\mathbf{K}\Lambda^{-1}. \tag{5.38}$$

On the other hand, we'll take $\left\langle \delta\mathfrak{r}\delta\mathfrak{r}^{\mathsf{T}}\right\rangle = \mathbf{D}t$ so that the same substitution gives

$$\mathbf{D}' = \Lambda\mathbf{D}\Lambda^{\mathsf{T}}. \tag{5.39}$$

We therefore need only provide the matrix $\Lambda$ to fully describe any gauge transformation. We will generally write $\Lambda$ as a block matrix (with each component a $3\times 3$ matrix), since $\delta\mathfrak{r} = (\boldsymbol{\delta r}, \boldsymbol{\delta\vartheta})$ is a pair of vectors (although this will break down when we look at the 1+5 representation in §5.2.6).

### 5.3.2 Body-centric and symmetric gauges

Note that throughout this section we will omit the $_{12}$ subscripts in all relative- and mis-translations, and relative- and mis-orientations, in favor of subscripts $_{(1)}$, $_{(2)}$, and $_{(s)}$ to indicate gauge (since absolute positions $\boldsymbol{R}_\alpha$ and orientations $\boldsymbol{\Theta}_\alpha$ are always in the global frame, their subscripts will retain their original meanings).

First, consider the 1- and 2-gauges in the 3+3 picture. We saw earlier that $\overline{\boldsymbol{\omega}}$ was invariant between these gauges, but we did find gauge dependence in the definitions

of $\boldsymbol{\delta\omega}$. Moreover, the relative translations $\boldsymbol{r}$ were dependent on the frame via the matrix $\mathbf{G}$. If we've measured $\overline{\boldsymbol{r}}_{(1)}$ in the 1-gauge, we can compute[7]

$$\overline{\boldsymbol{r}}_{(2)} = \boldsymbol{\Theta}_{\overline{2}}^{\mathsf{T}} \boldsymbol{\Theta}_{\overline{1}} \overline{\boldsymbol{r}} = \overline{\boldsymbol{\omega}}^{\mathsf{T}} \overline{\boldsymbol{r}}. \tag{5.40}$$

Now if we have the relative translation of a particular snapshot in the global reference frame, $(\boldsymbol{R}_1, \boldsymbol{\Theta}_1, \boldsymbol{R}_2, \boldsymbol{\Theta}_2)$, we can compute $\boldsymbol{r}_{(\alpha)} = \boldsymbol{\Theta}_\alpha^{\mathsf{T}} \boldsymbol{r}$ and $\boldsymbol{\delta r}_{(\alpha)} = \boldsymbol{r}_{(\alpha)} - \overline{\boldsymbol{r}}_{(\alpha)}$ for either frame $\alpha = 1$ or 2, or for the symmetric frame $\alpha = \mathrm{s}$ taking $\boldsymbol{\Theta}_\mathrm{s} = \boldsymbol{\Theta}_1 \omega^{1/2}$.

We would like to write a transformation directly between $\boldsymbol{\delta r}$ in these frames. Thus, we consider

$$\boldsymbol{R}_2 - \boldsymbol{R}_1 = \boldsymbol{\Theta}_1 \left( \overline{\boldsymbol{r}}_{(1)} + \boldsymbol{\delta r}_{(1)} \right) = \boldsymbol{\Theta}_2 \left( \overline{\boldsymbol{r}}_{(2)} + \boldsymbol{\delta r}_{(2)} \right). \tag{5.41}$$

Substituting $\overline{\boldsymbol{r}}_{(2)} = \overline{\boldsymbol{\omega}}^{\mathsf{T}} \overline{\boldsymbol{r}}_{(1)}$, left-multiplying by $\boldsymbol{\Theta}_2^{\mathsf{T}}$, and rearranging yields

$$\boldsymbol{\delta r}_{(2)} = \boldsymbol{\omega}^{\mathsf{T}} \left( \overline{\boldsymbol{r}}_{(1)} + \boldsymbol{\delta r}_{(1)} \right) - \overline{\boldsymbol{\omega}}^{\mathsf{T}} \overline{\boldsymbol{r}}_{(1)} = \overline{\boldsymbol{\omega}}^{\mathsf{T}} \left( \boldsymbol{\delta r}_{(1)} + \left( 1 - \boldsymbol{\delta\omega}_{(1)} \right) \overline{\boldsymbol{r}}_{(1)} \right). \tag{5.42}$$

Now for small misorientations we can write

$$\boldsymbol{\delta\omega}_{(1)} = 1 + [\boldsymbol{\delta\vartheta}_{(1)} \times], \tag{5.43}$$

so that the final transformation becomes

$$\boldsymbol{\delta r}_{(2)} = \overline{\boldsymbol{\omega}} \left( \boldsymbol{\delta r}_{(1)} + \overline{\boldsymbol{r}}_{(1)} \times \boldsymbol{\delta\vartheta}_{(1)} \right). \tag{5.44}$$

We see that the misorientation has coupled into the transformed mistranslation (we knew it had to be more than just a rotation of $\boldsymbol{\delta r}$ since this transformation can change the magnitude of $|\boldsymbol{\delta r}|$).

The inverse transformation is

$$\boldsymbol{\delta r}_{(1)} = \overline{\boldsymbol{\omega}}^{\mathsf{T}} \left( \boldsymbol{\delta r}_{(2)} - \overline{\boldsymbol{r}}_{(2)} \times \boldsymbol{\delta\vartheta}_{(2)} \right). \tag{5.45}$$

[7] This can be understood intuitively because $\overline{\boldsymbol{\omega}}$ rotates object 1 into object 2. Thus, $\overline{\boldsymbol{\omega}}^{\mathsf{T}}$ rotates from the the reference of 1 to the reference of 2 (i.e. the passive rotation picture).

Transforming into the symmetric gauge is slightly more problematic because the transformation involves $\boldsymbol{\omega}^{1/2} = \left(\overline{\boldsymbol{\omega}}\boldsymbol{\delta\omega}_{(2)}\right)^{1/2}$, which is difficult to deal with. We invoke a lemma[8] that for a rotation matrix $\boldsymbol{\Omega}$ and small vector $\boldsymbol{\theta}$,

$$\Big(\boldsymbol{\Omega}\big(1+[\boldsymbol{\theta}\times]\big)\Big)^{1/2} = \boldsymbol{\Omega}^{1/2}\Big(1+\Big[\big(1+\boldsymbol{\Omega}^{-1/2}\big)^{-1}\boldsymbol{\theta}\times\Big]\Big). \tag{5.46}$$

This allows repeating the above procedure for $\boldsymbol{\delta r}_{(\mathrm{s})}$ instead of $\boldsymbol{\delta r}_{(2)}$:

$$\boldsymbol{\delta r}_{(\mathrm{s})} = \overline{\boldsymbol{\omega}}^{-1/2}\Big(\boldsymbol{\delta r}_{(1)} + \big[\overline{\boldsymbol{r}}_{(1)}\times\big]\Big(1+\overline{\boldsymbol{\omega}}^{1/2}\big)^{-1}\boldsymbol{\delta\vartheta}_{(1)}\Big). \tag{5.47}$$

### 5.3.3 Rotated reference copies

Next we consider the effect of rotating the reference bodies. Suppose we apply the rotations $\Omega_\alpha$ to each reference body: $\boldsymbol{r}_\alpha^{(\mathrm{ref})\prime} = \boldsymbol{\Omega}_\alpha \boldsymbol{r}_\alpha^{(\mathrm{ref})}$. We then have $\boldsymbol{\Theta}'_\alpha = \boldsymbol{\Theta}_\alpha\boldsymbol{\Omega}_\alpha$. The surprising result here is the importance of the gauge choice in the transformation matrix. In the 1-gauge, we have

$$\overline{\boldsymbol{r}}'_{12} = \boldsymbol{\Omega}_1^{\mathsf{T}}\overline{\boldsymbol{r}}_{12} \tag{5.48}$$

$$\boldsymbol{\delta r}'_{12} = \boldsymbol{\Omega}_1^{\mathsf{T}}\boldsymbol{\delta r}_{12} \tag{5.49}$$

$$\boldsymbol{\delta\omega}'_{12} = \boldsymbol{\Omega}_1^{\mathsf{T}}\boldsymbol{\delta\omega}_{12}, \tag{5.50}$$

---

[8] This can be proved by squaring the right-hand side and applying the commutation relation

$$\boldsymbol{\Omega}\big[\boldsymbol{\theta}\times\big] = \big[\boldsymbol{\Omega}\boldsymbol{\theta}\times\big]\boldsymbol{\Omega}$$

to commute the $\boldsymbol{\Omega}^{1/2}$ left through $1+\big[(1+\boldsymbol{\Omega}^{-1/2})^{-1}\boldsymbol{\theta}\times\big]$, leaving

$$\boldsymbol{\Omega}\Big(1+\big[(1+\boldsymbol{\Omega}^{-1/2})(1+\boldsymbol{\Omega}^{-1/2})^{-1}\boldsymbol{\theta}\times\big]\Big),$$

which we can clearly see is equal to $\boldsymbol{\Omega}\big(1+[\boldsymbol{\theta}\times]\big)$.

while in the 2-gauge,

$$\overline{\boldsymbol{r}}'_{12} = \boldsymbol{\Omega}_2^\top \overline{\boldsymbol{r}}_{12} \tag{5.51}$$

$$\boldsymbol{\delta r}'_{12} = \boldsymbol{\Omega}_2^\top \boldsymbol{\delta r}_{12} \tag{5.52}$$

$$\boldsymbol{\delta \omega}'_{12} = \boldsymbol{\Omega}_2^\top \boldsymbol{\delta \omega}_{12}. \tag{5.53}$$

We see that rotations in the reference body 2 are entirely irrelevant to results in the 1-gauge, and vice versa. The symmetric gauge is more complicated, depending on both rotations in a nontrivial way in the general case that $\boldsymbol{\Omega}_1$, $\boldsymbol{\Omega}_2$, and $\overline{\boldsymbol{\omega}}_{12}$ do not commute

### 5.3.4 Reversing springs

Another gauge transformation we might like to make is to "reverse" a spring. That is, given $\mathfrak{Sr}_{12}$, what is $\mathfrak{Sr}_{21}$? In the symmetric gauge, this is as simple as reversing the signs of each vector. To reverse a spring in the 1-gauge (such that the result is in the *new* 1-gauge), we can use the fact that the 2-gauge of $\mathfrak{Sr}_{12}$ is equivalent to (that is, opposite in sign) to the 1-gauge of $\mathfrak{Sr}_{21}$, and vice versa, so we can first reverse the sign, then apply the proper gauge transformation. The symmetric gauge simply maps back onto its own negative.

### 5.3.5 3+3 to 1+5

The most complicated gauge transformation is transforming between the 3+3 and 1+5 pictures. Given a preferred displacement $\overline{\boldsymbol{r}}_{12}$, the preferred distance $\overline{d} = \left|\overline{\boldsymbol{r}}_{12}\right|$. Let $\hat{\boldsymbol{r}}_{12} = \overline{\boldsymbol{r}}_{12}/d$ be the unit vector corresponding to this inter-body axis. Then given a

configuration $\delta\mathbf{r}_{12}$ we can write

$$\delta d_{12} = \hat{\boldsymbol{r}}_{12} \cdot \boldsymbol{\delta r}_{12} \tag{5.54}$$

$$\delta \tau_{12} = \hat{\boldsymbol{r}}_{12} \cdot \boldsymbol{\delta \vartheta}_{12} \tag{5.55}$$

$$\boldsymbol{\delta \phi}_{12} = \mathsf{P}_{\perp \boldsymbol{r}_{12}} \left( \frac{\boldsymbol{\delta r}_{12}}{d} - \boldsymbol{\delta \vartheta}_{12} \right) \tag{5.56}$$

$$\boldsymbol{\delta \psi}_{12} = \frac{1}{2} \mathsf{P}_{\perp \boldsymbol{r}_{12}} \left( \frac{\boldsymbol{\delta r}_{12}}{d} + \boldsymbol{\delta \vartheta}_{12} \right) \tag{5.57}$$

where $\mathsf{P}_{\perp \boldsymbol{r}_{12}} = \mathbf{1} - \hat{\boldsymbol{r}}_{12} \otimes \hat{\boldsymbol{r}}_{12}$ is the projection perpendicular to $\hat{\boldsymbol{r}}_{12}$. Note that $\boldsymbol{\delta \phi}_{12}$ and $\boldsymbol{\delta \psi}_{12}$ contain between them only four degrees of freedom, but we have neglected to specify a decomposition. One simple workaround is to instead take just the magnitudes $\left|\boldsymbol{\delta \phi}_{12}\right|$ and $\left|\boldsymbol{\delta \psi}_{12}\right|$, reducing our configuration to four distinct dimensions, though we could also use cross products with $\hat{\boldsymbol{r}}_{12}$ to (somewhat arbitrarily) pick directions.

### 5.3.6 Large deviations

While not exactly a gauge transformation, it's worth considering the effect on our formalism of large deviations from equilibrium. This is important for numerical simulations of networks of generalized springs, particularly if there is orientational frustration, where the the product of all the preferred orientations around a loop is not necessarily the identity.

All the work so far has assumed small deviations, so that the misorientation $\boldsymbol{\delta\omega}$ behaved as a vector $\boldsymbol{\delta \vartheta}$. Here we will work out instead an analytical continuation of (5.1) that is valid for large misorientations as well[9].

[9] As far as the group structure is concerned, the misconfiguration is actually a *semidirect* product $\mathbb{R}^3 \rtimes \mathrm{SO}(3)$, rather than the direct product. The distinction is that in a direct product, both factors are normal subgroups of the product, so the group operation does not allow any mixing, while a semidirect product only requires one

We have two basic requirements for a good analytical continuation. The potential should be a continuous function of the relative coordinates $\mathbb{r}$, and it should be quadratic about its minimum, in such a way that we can relate our parameters to the stiffness tensor $\mathbf{K}$.

Žefran and Kumar [1] discuss a solution commonly used in robotics in which relative configurations are expressed in an exponential form, but this is unsatisfying because the resulting screw motion causes individual directions to lose their physical meanings.

We instead take a simpler approach in which the preferred rotation is allowed to be a function of the mistranslation. Suppose we have a spring with preferred configuration (in the 3+3 1-gauge) $\overline{\mathbb{r}} = \left(\overline{\boldsymbol{r}}, \overline{\boldsymbol{\omega}}\right)$ and actual configuration $\mathbb{r} = (\boldsymbol{r}, \boldsymbol{\omega})$. We define a rotation matrix $\boldsymbol{\omega}_\mathrm{r} = \mathsf{R}(\mathbf{W}\boldsymbol{\delta r})$, where $\boldsymbol{\delta r} = \boldsymbol{r} - \overline{\boldsymbol{r}}$ and $\mathbf{W}$ is an arbitrary matrix (we make no claims about (skew) symmetry). Now we define a new preferred relative orientation

$$\overline{\boldsymbol{\omega}}_\mathrm{r} = \boldsymbol{\omega}_\mathrm{r}\overline{\boldsymbol{\omega}} \tag{5.58}$$

and thus a new misorientation

$$\boldsymbol{\delta\omega}_\mathrm{r} = \boldsymbol{\omega}\overline{\boldsymbol{\omega}}_\mathrm{r}^\mathsf{T}. \tag{5.59}$$

From here, we can simply write the block-diagonal parts of (5.1),

$$\mathscr{F} = \boldsymbol{\delta r}^\mathsf{T}\mathbf{K}_\mathrm{r}\boldsymbol{\delta r} + \boldsymbol{\delta\theta}_\mathrm{r}^\mathsf{T}\mathbf{K}_\theta\boldsymbol{\delta\theta}_\mathrm{r}. \tag{5.60}$$

With $\mathbf{K}_\mathrm{r}$ and $\mathbf{K}_\theta$ symmetric $3\times 3$ matrices and $\mathbf{W}$ a general $3\times 3$ matrix, we again have 21 stiffness parameters. In the small misconfiguration limit, we find that

factor to be a normal subgroup, which ultimately allows mixing. In other words, if $(a_i, b_j) \in A \times B$ then $(a_1, b_1)(a_2, b_2) = (a_1 a_2, b_1 b_2)$, whereas if $(a_i, b_j) \in A \rtimes B$ then $(a_1, b_1)(a_2, b_2) = (a_1 \phi_{b_2}(a_2), b_1 b_2)$, where $\phi_b$ is some automorphism $\phi_b : A \to A$ for any $b \in B$. The upshot of this is that, while we can write quadratic forms in either $\mathbb{R}^3$ or SO(3), there is no best way to write a cross term between $\mathbb{R}^3$ and SO(3) that is linear in both at small deviations.

$\boldsymbol{\delta\vartheta}_{\mathrm{r}} = \boldsymbol{\delta\vartheta} - \mathbf{W}\boldsymbol{\delta r}$, so that

$$\mathscr{F} = \boldsymbol{\delta r}^{\mathsf{T}}\mathbf{K}_{\mathrm{r}}\boldsymbol{\delta r} + (\boldsymbol{\delta\vartheta} - \mathbf{W}\boldsymbol{\delta r})^{\mathsf{T}}\mathbf{K}_{\theta}(\boldsymbol{\delta\vartheta} - \mathbf{W}\boldsymbol{\delta r}). \tag{5.61}$$

Thus, the $6 \times 6$ stiffness tensor is

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}_{\mathrm{r}} + \mathbf{W}^{\mathsf{T}}\mathbf{K}_{\theta}\mathbf{W} & -\mathbf{W}^{\mathsf{T}}\mathbf{K}_{\theta} \\ -\mathbf{K}_{\theta}\mathbf{W} & \mathbf{K}_{\theta} \end{pmatrix}. \tag{5.62}$$

We can now use this equivalence to perform numerical minimizations and simulations

## 5.4 Composition laws

In this section we return to the small-deviation limit to write composition laws for the effective stiffness of a pair of springs connected in series or parallel.

Suppose we have three bodies (1, 2, and 3), connected by two generalized springs (A and B) in series. We would ultimately like to integrate out the position of the middle body, writing the interaction between the outside pair as a single effective spring:

$$\mathscr{F}(\mathbb{R}_1, \mathbb{R}_3) = \mathscr{F}_{\mathrm{A}}(\mathbb{R}_1, \mathbb{R}_2) + \mathscr{F}_{\mathrm{B}}(\mathbb{R}_2, \mathbb{R}_3). \tag{5.63}$$

### 5.4.1 Partition functions

To properly integrate out[10] the position of body 2, we must allow it to move freely, so that the partition function $\mathcal{Z}^{(\text{eff})} = \mathrm{e}^{-\beta\mathcal{F}^{(\text{eff})}}$

$$\mathcal{Z}^{(\text{eff})}(\mathbb{R}_1, \mathbb{R}_3) = \int \mathrm{d}\mathbb{R}_2\, \mathcal{Z}_{\text{A}}(\mathbb{R}_1, \mathbb{R}_2)\mathcal{Z}_{\text{B}}(\mathbb{R}_2, \mathbb{R}_3). \tag{5.64}$$

Note that the units of $\mathcal{Z}^{(\text{eff})}$ are different from the units of $\mathcal{Z}_{\text{A,B}}$. This is not a problem, since we're interested in $\log\mathcal{Z}$, so any unit conversions come out as a constant shift. Since the free energies (and hence partition functions) are defined in terms of the relative coordinates, we prefer to write our integral in these coordinates, integrating over the relative coordinates of both springs $\mathbb{r}_{\text{A,B}}$, and multiplying by a $\delta$-function to ensure that $\mathbb{r}_{\text{B}} \circ \mathbb{r}_{\text{A}} = \mathbb{r} = \mathbb{R}_3 - \mathbb{R}_1$. Finally, to integrate over the rotations we assume small misorientations. We take preferred configurations $\overline{\mathbb{r}}_{\text{A,B}} = \left(\overline{\boldsymbol{r}}_{\text{A,B}}, \overline{\boldsymbol{\theta}}_{\text{A,B}}\right)$ and misconfigurations $\delta\mathbb{r}_{\text{A,B}} = \left(\boldsymbol{\delta r}_{\text{A,B}}, \boldsymbol{\delta\vartheta}_{\text{A,B}}\right)$. Then

$$\begin{aligned}\mathcal{Z}^{(\text{eff})}(\mathbb{r}) = \int \mathrm{d}^3\boldsymbol{\delta r}_{\text{A}}\, \mathrm{d}^3\boldsymbol{\delta\vartheta}_{\text{A}}\, \mathrm{d}^3\boldsymbol{\delta r}_{\text{B}}\, \mathrm{d}^3\boldsymbol{\delta\vartheta}_{\text{B}} \\ \mathrm{e}^{-\beta(\mathcal{F}_{\text{A}}+\mathcal{F}_{\text{B}})}\delta^6\left((\mathbb{r}_{\text{B}} \circ \mathbb{r}_{\text{A}}) - \mathbb{r}\right).\end{aligned} \tag{5.65}$$

We need to write $\delta^6\left((\mathbb{r}_{\text{B}} \circ \mathbb{r}_{\text{A}}) - \mathbb{r}\right)$ in terms of our integration variables $\boldsymbol{\delta r}_{\text{A}}$, $\boldsymbol{\delta r}_{\text{B}}$, $\boldsymbol{\delta\vartheta}_{\text{A}}$, and $\boldsymbol{\delta\vartheta}_{\text{B}}$, and our free variables $\boldsymbol{\delta r}$ and $\boldsymbol{\delta\vartheta}$. Once we've done this, the $\delta$-function can be converted into an exponential and the whole thing integrated to yield a new Gaussian for $\mathcal{Z}^{(\text{eff})}(\mathbb{r})$, which will give the effective combined free energy.

[10] We could take a shortcut instead by simply determining the minimum-energy position for body 2 and take that as the free energy, i.e.

$$\mathcal{F}^{(\text{eff})}(\mathbb{R}_1, \mathbb{R}_3) = \min_{\mathbb{R}_2} \mathcal{F}(\mathbb{R}_1, \mathbb{R}_2, \mathbb{R}_3).$$

This should be similar, but off by an order-constant factor. But this minimization is difficult and may not be tractable analytically.

### 5.4.2 Constraints

This six-dimensional $\delta$-function, $\delta^6\big((\mathbb{r}_B \circ \mathbb{r}_A) - \mathbb{r}\big)$ has a translational and an orientational part, but it will take some work to write these terms explicitly. In order to do so, we must write the coordinates of the total relative configuration $\mathbb{r}_B \circ \mathbb{r}_A$ in some gauge, which means we must first decide on a gauge. While it's simplest to work in the 1-gauge, the result is (predictably) not symmetric in the springs A and B (this derivation is left as an exercise to the reader). We will therefore work instead in the symmetric gauge (and all relative coordinates are assumed to be in this gauge). Once we've written the total relative configuration, we can break this up into a large constant part (which is the preferred relative configuration $\overline{\mathbb{r}}$) and a small perturbation $\delta\mathbb{r}$, which we must write in terms of $\delta\mathbb{r}_A$ and $\delta\mathbb{r}_B$.

### 5.4.3 Orientational constraint

We begin with orientation. The total relative orientation $\boldsymbol{\omega} = \boldsymbol{\omega}_A\boldsymbol{\omega}_B$ can be expanded in either of the following ways:

$$\boldsymbol{\omega} = \overline{\boldsymbol{\omega}}^{1/2}\big(1 + [\boldsymbol{\delta\vartheta}\times]\big)\overline{\boldsymbol{\omega}}^{1/2} \tag{5.66}$$

$$\boldsymbol{\omega} = \boldsymbol{\omega}_A\boldsymbol{\omega}_B = \overline{\boldsymbol{\omega}}_A^{1/2}\big(1 + [\boldsymbol{\delta\vartheta}_A\times]\big)\overline{\boldsymbol{\omega}}_A^{1/2}\overline{\boldsymbol{\omega}}_B^{1/2}\big(1 + [\boldsymbol{\delta\vartheta}_B\times]\big)\overline{\boldsymbol{\omega}}_B^{1/2}. \tag{5.67}$$

Commuting all the large rotations to the left yields

$$\overline{\boldsymbol{\omega}}\Big(1 + \Big[\overline{\boldsymbol{\omega}}^{-1/2}\boldsymbol{\delta\vartheta}\times\Big]\Big) = \overline{\boldsymbol{\omega}}_A\overline{\boldsymbol{\omega}}_B\Big(1 + \Big[\overline{\boldsymbol{\omega}}_B^{-1}\overline{\boldsymbol{\omega}}_A^{-1/2}\boldsymbol{\delta\vartheta}_A\times\Big] + \Big[\overline{\boldsymbol{\omega}}_B^{-1/2}\boldsymbol{\delta\vartheta}_B\times\Big]\Big). \tag{5.68}$$

We can divide out the $\overline{\boldsymbol{\omega}} = \overline{\boldsymbol{\omega}}_A\overline{\boldsymbol{\omega}}_B$ (this is actually the definition of $\overline{\boldsymbol{\omega}}$) and switch to a vector picture since all that's left is small skew-symmetric matrices. Noting that $\overline{\boldsymbol{\omega}}_B^{-1}\overline{\boldsymbol{\omega}}_A^{-1/2} = \overline{\boldsymbol{\omega}}^{-1}\overline{\boldsymbol{\omega}}_A^{1/2}$ allows us to write finally the constraint $\delta^3_{\text{orient}}$,

$$\boldsymbol{\delta\vartheta} = \overline{\boldsymbol{\omega}}^{-1/2}\overline{\boldsymbol{\omega}}_A^{1/2}\boldsymbol{\delta\vartheta}_A + \overline{\boldsymbol{\omega}}^{1/2}\overline{\boldsymbol{\omega}}_B^{-1/2}\boldsymbol{\delta\vartheta}_B = \overline{\boldsymbol{\zeta}}_{+A}^{\mathsf{T}}\boldsymbol{\delta\vartheta}_A + \overline{\boldsymbol{\zeta}}_{-B}^{\mathsf{T}}\boldsymbol{\delta\vartheta}_B, \tag{5.69}$$

where we have defined

$$\overline{\boldsymbol{\zeta}}_{\pm\alpha} \equiv \overline{\boldsymbol{\omega}}_{\alpha}^{\mp 1/2} \overline{\boldsymbol{\omega}}^{\pm 1/2}, \tag{5.70}$$

the transformations between the symmetric gauges of each spring and the symmetric gauge of the composed spring.

### 5.4.4 Translational constraint

The mistranslation constraints are worked out similarly, starting with

$$\boldsymbol{R}_2 - \boldsymbol{R}_1 = \boldsymbol{\Theta}_1 \boldsymbol{\omega}_{\mathrm{A}}^{1/2} \big( \overline{\boldsymbol{r}}_{\mathrm{A}} + \boldsymbol{\delta r}_{\mathrm{A}} \big) \tag{5.71}$$

$$\boldsymbol{R}_3 - \boldsymbol{R}_2 = \boldsymbol{\Theta}_3 \boldsymbol{\omega}_{\mathrm{B}}^{-1/2} \big( \overline{\boldsymbol{r}}_{\mathrm{B}} + \boldsymbol{\delta r}_{\mathrm{B}} \big) \tag{5.72}$$

$$\boldsymbol{R}_3 - \boldsymbol{R}_1 = \boldsymbol{\Theta}_1 \boldsymbol{\omega}^{1/2} \big( \overline{\boldsymbol{r}} + \boldsymbol{\delta r} \big). \tag{5.73}$$

We note that $\boldsymbol{\Theta}_3 = \boldsymbol{\Theta}_1 \boldsymbol{\omega}$, so that the $\boldsymbol{\Theta}_1$ can be divided out, leaving

$$\overline{\boldsymbol{r}} + \boldsymbol{\delta r} = \boldsymbol{\zeta}_{+\mathrm{A}}^{\mathsf{T}} \big( \overline{\boldsymbol{r}}_{\mathrm{A}} + \boldsymbol{\delta r}_{\mathrm{A}} \big) + \boldsymbol{\zeta}_{-\mathrm{B}}^{\mathsf{T}} \big( \overline{\boldsymbol{r}}_{\mathrm{A}} + \boldsymbol{\delta r}_{\mathrm{A}} \big). \tag{5.74}$$

This leaves us with the question of how to write $\boldsymbol{\zeta}_{\pm\mu}$ in terms of $\overline{\boldsymbol{\zeta}}_{\pm\mu}$. We can apply the lemma (5.46) to expand $\boldsymbol{\omega}^{1/2}$ from (5.66):

$$\boldsymbol{\omega}^{1/2} = \Big( 1 + \big[ \mathbf{Y}_{-} \boldsymbol{\delta\vartheta} \times \big] \Big) \overline{\boldsymbol{\omega}}^{1/2} = \overline{\boldsymbol{\omega}}^{1/2} \Big( 1 + \big[ \mathbf{Y}_{+} \boldsymbol{\delta\vartheta} \times \big] \Big), \tag{5.75}$$

where we've defined

$$\mathbf{Y}_{\pm} = \Big( 1 + \overline{\boldsymbol{\omega}}^{\pm 1/2} \Big)^{-1} = \mathbf{Y}_{\mp}^{\mathsf{T}} \tag{5.76}$$

for page economy. Thus, we find

$$\boldsymbol{\zeta}_{\pm\mu} = \Big( 1 \mp \big[ \mathbf{Y}_{\pm\mu} \boldsymbol{\delta\vartheta}_{\mu} \times \big] \pm \big[ \overline{\boldsymbol{\omega}}_{\mu}^{\mp 1/2} \mathbf{Y}_{\mp} \boldsymbol{\delta\vartheta} \times \big] \Big) \overline{\boldsymbol{\zeta}}_{\pm\mu}. \tag{5.77}$$

We can now rewrite (5.74), separating out the large part,

$$\overline{\boldsymbol{r}} = \overline{\boldsymbol{\zeta}}_{+\mathrm{A}}^{\mathsf{T}} \overline{\boldsymbol{r}}_{\mathrm{A}} + \overline{\boldsymbol{\zeta}}_{-\mathrm{B}}^{\mathsf{T}} \overline{\boldsymbol{r}}_{\mathrm{B}}, \tag{5.78}$$

as a definition of $\overline{\boldsymbol{r}}$, and taking the small part as the constraint:

$$\begin{aligned}\boldsymbol{\delta r} =& \overline{\boldsymbol{\zeta}}_{+\mathrm{A}}^{\mathsf{T}}\Big(\boldsymbol{\delta r}_{\mathrm{A}} - \big[\overline{\boldsymbol{r}}_{\mathrm{A}}\times\big]\mathbf{Y}_{+\mathrm{A}}\boldsymbol{\delta\vartheta}_{\mathrm{A}} + \big[\overline{\boldsymbol{r}}_{\mathrm{A}}\times\big]\overline{\boldsymbol{\omega}}_{\mathrm{A}}^{-1/2}\mathbf{Y}_{-}\boldsymbol{\delta\vartheta}\Big)\\ &+ \overline{\boldsymbol{\zeta}}_{-\mathrm{B}}^{\mathsf{T}}\Big(\boldsymbol{\delta r}_{\mathrm{B}} + \big[\overline{\boldsymbol{r}}_{\mathrm{B}}\times\big]\mathbf{Y}_{-\mathrm{B}}\boldsymbol{\delta\vartheta}_{\mathrm{B}} - \big[\overline{\boldsymbol{r}}_{\mathrm{B}}\times\big]\overline{\boldsymbol{\omega}}_{\mathrm{B}}^{1/2}\mathbf{Y}_{+}\boldsymbol{\delta\vartheta}\Big).\end{aligned} \tag{5.79}$$

### 5.4.5 Putting it all together

We can now substitute (5.69) and (5.79) into the $\delta$-function in (5.65), rewriting $\delta^3(\boldsymbol{x}) = \int \mathrm{d}^3\boldsymbol{k}\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}$, and introducing two new vector integration variables, $\boldsymbol{k}$ conjugate to the translational part and $\boldsymbol{q}$ conjugate to the orientational part. This leaves

$$\mathcal{Z}^{(\mathrm{eff})} = \int \mathrm{d}^{18}\boldsymbol{x}\,\mathrm{e}^{-\frac{1}{2}\boldsymbol{x}^{\mathsf{T}}\mathbf{A}\boldsymbol{x}-\boldsymbol{b}\cdot\boldsymbol{x}} \propto \mathrm{e}^{\frac{1}{2}\boldsymbol{b}^{\mathsf{T}}\mathbf{A}^{-1}\boldsymbol{b}}, \tag{5.80}$$

where $\boldsymbol{x} = \big(\boldsymbol{\delta r}_{\mathrm{A}}, \boldsymbol{\delta\vartheta}_{\mathrm{A}}, \boldsymbol{\delta r}_{\mathrm{B}}, \boldsymbol{\delta\vartheta}_{\mathrm{B}}, \boldsymbol{k}, \boldsymbol{q}\big)$ is an 18-dimensional vector made of six 3-vectors, and $\mathbf{A}$ is given by

$$\mathbf{A} = \begin{pmatrix} \mathbf{K}_{\mathrm{A}}^{\mathrm{rr}} & \mathbf{K}_{\mathrm{A}}^{\mathrm{r}\theta} & 0 & 0 & \mathrm{i}\overline{\boldsymbol{\zeta}}_{+\mathrm{A}} & 0 \\ \mathbf{K}_{\mathrm{A}}^{\theta\mathrm{r}} & \mathbf{K}_{\mathrm{A}}^{\theta\theta} & 0 & 0 & \mathrm{i}\mathbf{Z}_{+\mathrm{A}}\overline{\boldsymbol{\zeta}}_{+\mathrm{A}} & \mathrm{i}\overline{\boldsymbol{\zeta}}_{+\mathrm{A}} \\ 0 & 0 & \mathbf{K}_{\mathrm{B}}^{\mathrm{rr}} & \mathbf{K}_{\mathrm{B}}^{\mathrm{r}\theta} & \mathrm{i}\overline{\boldsymbol{\zeta}}_{-\mathrm{B}} & 0 \\ 0 & 0 & \mathbf{K}_{\mathrm{B}}^{\theta\mathrm{r}} & \mathbf{K}_{\mathrm{B}}^{\theta\theta} & \mathrm{i}\mathbf{Z}_{-\mathrm{B}}\overline{\boldsymbol{\zeta}}_{-\mathrm{B}} & \mathrm{i}\overline{\boldsymbol{\zeta}}_{-\mathrm{B}} \\ \mathrm{i}\overline{\boldsymbol{\zeta}}_{+\mathrm{A}}^{\mathsf{T}} & \mathrm{i}\overline{\boldsymbol{\zeta}}_{+\mathrm{A}}^{\mathsf{T}}\mathbf{Z}_{+\mathrm{A}}^{\mathsf{T}} & \mathrm{i}\overline{\boldsymbol{\zeta}}_{-\mathrm{B}}^{\mathsf{T}} & \mathrm{i}\overline{\boldsymbol{\zeta}}_{-\mathrm{B}}^{\mathsf{T}}\mathbf{Z}_{-\mathrm{B}}^{\mathsf{T}} & 0 & 0 \\ 0 & \mathrm{i}\overline{\boldsymbol{\zeta}}_{+\mathrm{A}}^{\mathsf{T}} & 0 & \mathrm{i}\overline{\boldsymbol{\zeta}}_{-\mathrm{B}}^{\mathsf{T}} & 0 & 0 \end{pmatrix}, \tag{5.81}$$

with $\mathbf{Z}_{\pm\mu} = \pm\mathbf{Y}_{\mp\mu}[\overline{\boldsymbol{r}}_{\mathrm{A}}\times]$. The $6\cdot3$-vector $\boldsymbol{b}$ has only two of its six 3-vectors nonzero:

$$\begin{pmatrix} \boldsymbol{b}_{\boldsymbol{k}} \\ \boldsymbol{b}_{\boldsymbol{q}} \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{X} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \mathrm{i}\boldsymbol{\delta r} \\ \mathrm{i}\boldsymbol{\delta\vartheta} \end{pmatrix} \tag{5.82}$$

$$\boldsymbol{b}_{\boldsymbol{\delta r}_{\mathrm{A}}} = \boldsymbol{b}_{\boldsymbol{\delta\vartheta}_{\mathrm{A}}} = \boldsymbol{b}_{\boldsymbol{\delta r}_{\mathrm{B}}} = \boldsymbol{b}_{\boldsymbol{\delta\vartheta}_{\mathrm{B}}} = 0, \tag{5.83}$$

where $\mathbf{X} = [\overline{\boldsymbol{r}}_{\mathrm{B}}\times]\overline{\boldsymbol{\omega}}_{\mathrm{B}}^{1/2}\mathbf{Y}_{+} - [\overline{\boldsymbol{r}}_{\mathrm{A}}\times]\overline{\boldsymbol{\omega}}_{\mathrm{A}}^{-1/2}\mathbf{Y}_{-}$.

As we see in (5.80) we need to invert $\mathbf{A}$. For a general block-matrix, we have

$$\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}^{-1} = \begin{pmatrix} \left(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\right)^{-1} & \left(\mathbf{A}_{21} - \mathbf{A}_{22}\mathbf{A}_{12}^{-1}\mathbf{A}_{11}\right)^{-1} \\ \left(\mathbf{A}_{12} - \mathbf{A}_{11}\mathbf{A}_{21}^{-1}\mathbf{A}_{22}\right)^{-1} & \left(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\right)^{-1} \end{pmatrix} \tag{5.84}$$

If we take '2' in (5.84) to represent the nonzero directions in $\boldsymbol{b}$ (that is, $\boldsymbol{k}$ and $\boldsymbol{q}$) then we need only compute $\left(\mathbf{A}^{-1}\right)_{22}$. Moreover, since $\mathbf{A}_{22} = 0$, we have simply $(\mathbf{A}_{22})^{-1} = -\left(\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\right)^{-1}$. Further, $\mathbf{A}_{11}^{-1}$ is the block matrix of the inverse stiffness matrices $\mathbf{Q} = \mathbf{K}^{-1}$:

$$\mathbf{A}_{11}^{-1} = \begin{pmatrix} \mathbf{Q}_{\mathrm{A}}^{\mathrm{rr}} & \mathbf{Q}_{\mathrm{A}}^{\mathrm{r}\theta} & 0 & 0 \\ \mathbf{Q}_{\mathrm{A}}^{\theta\mathrm{r}} & \mathbf{Q}_{\mathrm{A}}^{\theta\theta} & 0 & 0 \\ 0 & 0 & \mathbf{Q}_{\mathrm{B}}^{\mathrm{rr}} & \mathbf{Q}_{\mathrm{B}}^{\mathrm{r}\theta} \\ 0 & 0 & \mathbf{Q}_{\mathrm{B}}^{\theta\mathrm{r}} & \mathbf{Q}_{\mathrm{B}}^{\theta\theta} \end{pmatrix}. \tag{5.85}$$

Thus we can multiply out $\mathscr{F} = \boldsymbol{b}^{\mathsf{T}}\mathbf{A}^{-1}\boldsymbol{b} = -\frac{1}{2}\delta\mathfrak{r}^{\mathsf{T}}\mathbf{K}\delta\mathfrak{r}$ to get the stiffness of the composed spring,

$$\mathbf{K}_{\text{series}} = \mathbf{X}^{\triangle\mathsf{T}}\left(\overline{\boldsymbol{\zeta}}_{+\mathrm{A}}^{\mathsf{T}}\mathbf{Z}_{+\mathrm{A}}^{\nabla\mathsf{T}}\mathbf{Q}_{\mathrm{A}}\mathbf{Z}_{+\mathrm{A}}^{\nabla}\overline{\boldsymbol{\zeta}}_{+\mathrm{A}} + \overline{\boldsymbol{\zeta}}_{-\mathrm{B}}^{\mathsf{T}}\mathbf{Z}_{-\mathrm{B}}^{\nabla\mathsf{T}}\mathbf{Q}_{\mathrm{B}}\mathbf{Z}_{-\mathrm{B}}^{\nabla}\overline{\boldsymbol{\zeta}}_{-\mathrm{B}}\right)^{-1}\mathbf{X}^{\triangle}, \tag{5.86}$$

where we define the notation

$$\mathbf{M}^{\triangle} \equiv \begin{pmatrix} 1 & \mathbf{M} \\ 0 & 1 \end{pmatrix}, \qquad \mathbf{M}^{\nabla} \equiv \begin{pmatrix} 1 & 0 \\ \mathbf{M} & 1 \end{pmatrix}. \tag{5.87}$$

### 5.4.6 Parallel composition

We can also compose springs in parallel, though it's a bit anticlimactic after the preceding derivation. If two springs have the same preferred relative configuration $\overline{\mathfrak{r}}$, then the stiffnesses just add normally: $\mathbf{K}_{\text{parallel}} = \mathbf{K}_{\mathrm{A}} + \mathbf{K}_{\mathrm{B}}$.

If, on the other hand, the preferred displacements or orientations differ, then we are no longer in a small misconfiguration regime, so the approximations we relied on

in the series calculation no longer hold and we therefore cannot construct an analogue.

## BIBLIOGRAPHY

[1] M. Žefran and V. Kumar, "A geometrical approach to the study of the Cartesian stiffness matrix", *J. Mech. Design* **124**, 30 (2002).

CHAPTER 6

# EXTRACTING ELASTIC PARAMETERS FROM MD

This chapter is roughly the same as our recently published PRE rapid communication [1], with the addition of Figures 6.2 and 6.3, and some extra sectioning and description in §6.4.

## 6.1 Introduction

Large protein assemblies are pertinent to most of the soft-matter physics in cells; how can one calculate their elastic properties and corresponding dynamics? Such assemblies are too large to handle by all-atom simulations, but numerical coarse graining techniques are opening the door to direct simulations [2]. An ideal coarse-graining would involve rather simple parametrizations for the purposes of human understanding, analytic treatment, transmission to other researchers, and building up coarse-grained models [3, 4, 5]. In this chapter we propose a new approach to extract such simplified parameters from all-atom MD simulations of small subsystems. Our approach should generalize well to any system in which the proteins are primarily structural. Moreover, because our simulations involve only a few proteins, they are tractable even without supercomputers.

Our program is to break up a large assembly as a network of many discrete units connected by several kinds of (typically) pairwise interactions modeled as generalized springs. We treat each unit—typically a single protein or domain—as a rigid body, which thus has only six degrees of freedom. We approach each interaction, *one kind at a time*, by simulating just the pair of interacting units and measuring the trajectory $\boldsymbol{x}(t)$ of the positions and orientations of each unit. Our aim is, from these observed trajectories, to extract the parameters for an effective Hamiltonian

and equation of motion for each spring, and then reassemble the springs into the coarse-grained network.

We model $\boldsymbol{x}(t)$ as an overdamped random walk in a (biased) harmonic potential. This walk is parametrized primarily by two important tensors: one to describe the shape of the harmonic well, and the other to describe the (mainly hydrodynamic) damping and the associated stochastic noise. Combining these tensors gives a matrix whose eigenvalues are the relaxation rates. Thus by measuring short-time dynamic correlation functions together with the position distribution, we can identify whether the simulation has sufficiently sampled the equilibrium ensemble during the simulation time, or in case of an overall drift, we can compute the external forces needed to shift the equilibrium to the biologically proper configuration. This is similar in spirit to computing a potential of mean force or free energy landscape with Jarzynski's equality[6, 7], except that our coarse-grained $\boldsymbol{x}$ has more than one component, and (at minimum) represents angular degrees of freedom in addition to stretching. As an application, we simulate several inter-domain interactions in the HIV capsid and estimate the Young's modulus (which can be measured experimentally) and Poisson ratio of the capsid lattice, as well as the relaxation rate of the breathing mode.

## 6.2 Coarse grained stochastic dynamics

We represent our system as a vector of generalized coordinates $x_i$, $i = 1 \ldots N$, where $N$ is far smaller than the number of atoms and is obtained by some form of coarse-graining. Our objective is to parametrize and determine from simulation (i) an effective free energy potential function $\mathscr{F}(\boldsymbol{x})$, and (ii) an equation of motion, for the coarse-grained coordinates.

We assume the coarse-grained degrees of freedom are overdamped: this is true at

time scales much longer than the "ballistic scale" of local bond vibrations ($t_{\rm bal} \sim \Gamma m \sim$ 1ps, where $m \sim$ 10kDa is the mass of a protein). Then the dynamics is a continuous-time random walk:

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{\Gamma}\boldsymbol{f}(\boldsymbol{x},t) + \boldsymbol{\zeta}(t), \tag{6.1}$$

where $\boldsymbol{\Gamma}$ is the (symmetric) *mobility tensor*, $\boldsymbol{f}(\boldsymbol{x},t)$ is the force, $\boldsymbol{\zeta}(t)$ is a (Gaussian) stochastic function satisfying

$$\left\langle \boldsymbol{\zeta}(t) \otimes \boldsymbol{\zeta}(t') \right\rangle = 2\mathbf{D}\delta(t-t'), \tag{6.2}$$

and $\mathbf{D}$ is the *diffusion tensor*. For detailed balance, $\mathbf{D} = k_{\rm B}T\boldsymbol{\Gamma}$ at temperature $T$. We can expand the potential to second order about a point $\overline{\boldsymbol{x}}$,

$$\mathscr{F}(\boldsymbol{x}) = \mathscr{F}_0 - \overline{\boldsymbol{f}}\cdot(\boldsymbol{x}-\overline{\boldsymbol{x}}) + \frac{1}{2}(\boldsymbol{x}-\overline{\boldsymbol{x}})\mathbf{K}(\boldsymbol{x}-\overline{\boldsymbol{x}}), \tag{6.3}$$

where $\mathbf{K}$ is the (symmetric) *stiffness tensor*; then the force in (6.1) is $\boldsymbol{f}(\boldsymbol{x}) = \overline{\boldsymbol{f}} - \mathbf{K}(\boldsymbol{x}-\overline{\boldsymbol{x}})$. From measuring coordinate covariances in the simulation, we obtain $\mathbf{K}$:

$$\mathbf{G} \equiv \left\langle [\boldsymbol{x}-\overline{\boldsymbol{x}}] \otimes [\boldsymbol{x}-\overline{\boldsymbol{x}}] \right\rangle = k_{\rm B}T\mathbf{K}^{-1}. \tag{6.4}$$

If the static effective potential were our only interest, and if our runs were always long enough to equilibrate our system, there would have been no need to model the dynamics (6.1). As we do need the dynamics, we determine the diffusion tensor $\mathbf{D}$ (and hence $\boldsymbol{\Gamma}$) by measuring the correlation function at short times between the ballistic and relaxation time scales (see below) during which the deterministic term in (6.1) is less important than the noise:

$$\mathbf{D} = \frac{\left\langle [\boldsymbol{x}(t')-\boldsymbol{x}(t)] \otimes [\boldsymbol{x}(t')-\boldsymbol{x}(t)] \right\rangle}{2|t'-t|} \equiv \frac{\mathbf{W}(t'-t)}{2|t'-t|}. \tag{6.5}$$

We calculate $\mathbf{D}$ by fitting $\mathbf{W}(\Delta t)$ to a line over offsets $t_{\rm bal} \ll \Delta t \ll \tau$, the relaxation times, weighting each point by $\sigma^2_{\mathbf{W}}(\Delta t) \propto (\Delta t)^3$. Notice that since $\boldsymbol{\Gamma}$ pertains to short-

time dynamics, it is correctly measured even in runs too short to equilibrate in the potential well.

If we transform into coordinates $\widetilde{\boldsymbol{x}} \equiv \boldsymbol{\Gamma}^{-1/2}\boldsymbol{x}$ then the equation of motion becomes

$$\frac{\mathrm{d}\widetilde{\boldsymbol{x}}}{\mathrm{d}t} = \boldsymbol{\Gamma}^{1/2}\overline{\boldsymbol{f}} - \mathbf{R}\big(\widetilde{\boldsymbol{x}} - \overline{\widetilde{\boldsymbol{x}}}\big) + \widetilde{\boldsymbol{\zeta}}(t), \tag{6.6}$$

where

$$\big\langle \widetilde{\zeta}_\alpha(t)\widetilde{\zeta}_\beta(t')\big\rangle = 2k_\mathrm{B}T\delta_{\alpha\beta}\delta(t-t'), \tag{6.7}$$

and the *relaxation matrix* $\mathbf{R} = \boldsymbol{\Gamma}^{1/2}\mathbf{K}\boldsymbol{\Gamma}^{1/2}$ (which has units [time]$^{-1}$) is simply the stiffness tensor in our transformed frame. The eigenvalues of $\mathbf{R}$ are the decay rates $\tau_\alpha^{-1}$ for the relaxation normal modes $\alpha$.

The correlation time for a mode is the same as its relaxation time, so the relative error in $\mathbf{K}$ for mode $\alpha$ is of order $\sqrt{\tau_\alpha/\tau_\mathrm{run}}$, where $\tau_\mathrm{run}$ is the total run time. Thus, if all the $\tau_\alpha \ll \tau_\mathrm{run}$, our estimate (6.4) of $\mathbf{K}$ is valid. But if $\tau_\alpha \sim \tau_\mathrm{run}$ for some direction, not only are errors large, but the initial deviation may still be relaxing over the entire run, which is often visible as a steady drift of the coordinates with mean velocity $\overline{\boldsymbol{v}}$. Averaging over time gives a large spurious variance in the drifting directions, leading to an underestimate of the corresponding stiffness.

## 6.3 Application to HIV capsid

The elastic and dynamic properties of viruses in general are of particular importance in understanding the mechanisms by which they assemble and disassemble. The assembly must be reliable enough to produce capsids capable of surviving the harsh intercellular environment, while still being able to disassemble upon entering a new host cell. HIV in particular is unique because of its characteristic conical capsids [8], whose mechanism of formation is yet unsettled.

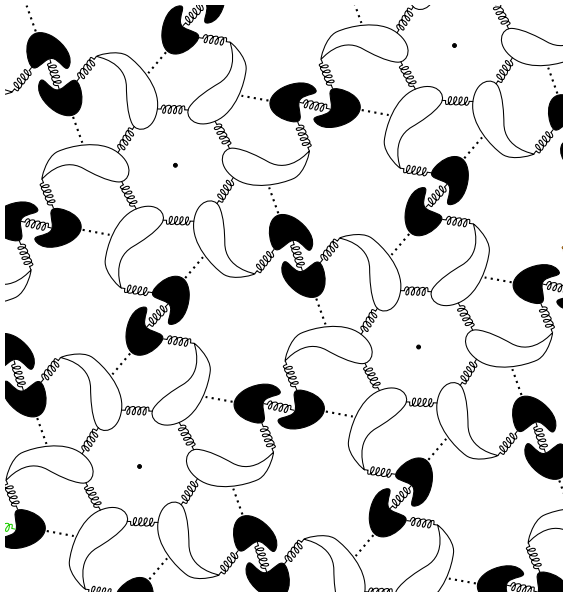
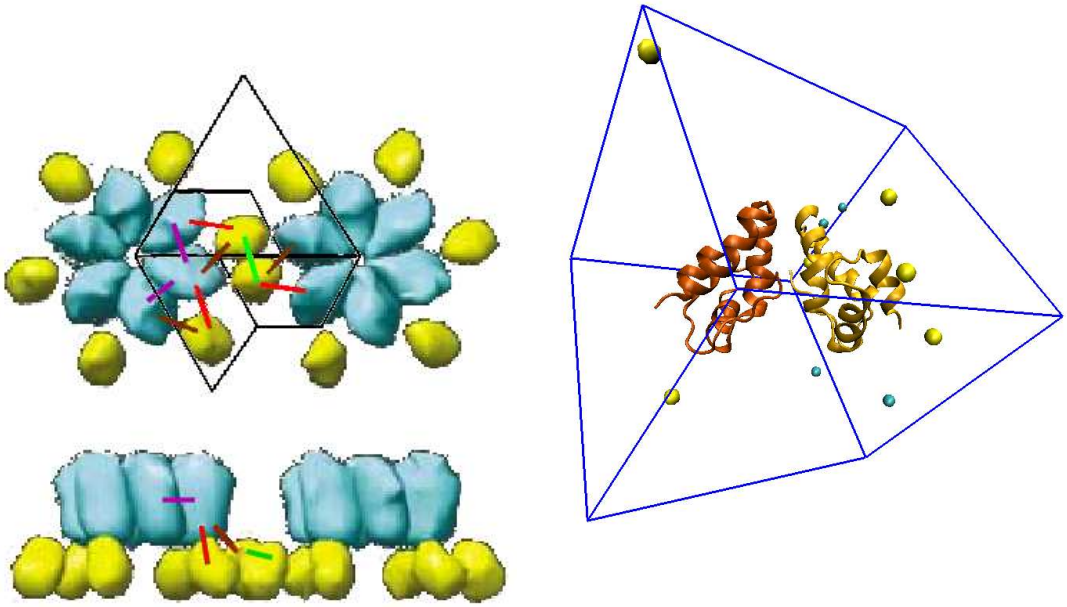


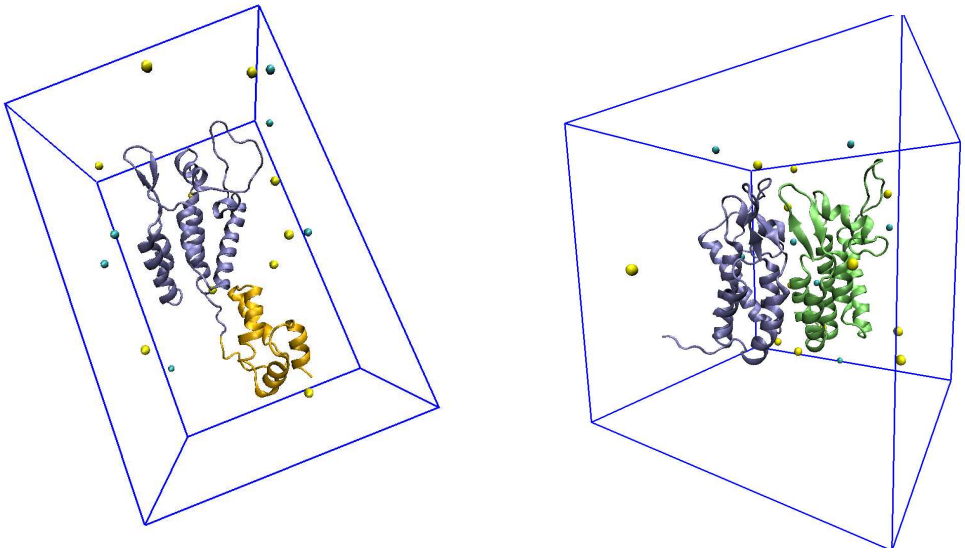
Figure 6.1: Diagram of interactions in the HIV capsid lattice. The black and white shapes represent the dimer-forming CTD and the hexamer-forming NTD, respectively. Springs represent the three different bonds we are interested in, and dotted lines represent the fourth bond we are ignoring.

The HIV capsid protein (CA) consists of two globular domains: the larger 145-amino acid N-terminal domain (NTD) has a radius 1.3nm and the smaller 70-amino acid C-terminal domain (CTD) has a radius 1.7nm; we treat these as two separate units. The NTD and CTD are connected covalently by a flexible linker; there is also an NTD–NTD interaction (which forms hexamers in the capsid structure), a CTD–CTD interaction (which forms symmetric dimers in the structure), and an NTD–CTD interaction between neighboring proteins around a hexamer. These four interactions are shown schematically in Figure 6.1 and a cryo-EM reconstruction is reproduced in Figure 6.2(a). We believe the NTD–CTD interaction to be the weakest, and the known structure is also poorest, so we will ignore it from now on. We therefore simulate each other pair in isolation, using structures from the Protein Data Bank[1].

[1]For full-length protein ("linker") we use cryo-EM structure 3DIK [10]; for NTD we use the NMR structure 1GWP [11] fitted to the homologous MLV hexamer crystal structure 1U7K [12]; and for CTD we use the crystal structure 1AUM [13].



(a) (b)



(c) (d)

Figure 6.2: Models of mature HIV. In all subfigures, the NTD is represented with blue/green colors, while the CTD is represented with yellow/red colors. (a) Electron microscopy reconstruction of a piece of mature HIV lattice, from Wright et al. [9]. Above, top view; Below, side view. The different bonds are drawn in analogue to Figure 4.2 and for comparison with Figure 6.1. (b–d) Simulation cells for (b) CTD–CTD dimer, (c) CTD–NTD linker, and (d) NTD–NTD heterodimer.

We carried out our simulations using a modified version of the NAMD [14] package with the CHARMM22 force field [15]. Our proteins are in a periodic cell 5 to 9nm to a side using the TIP3P model for explicit water and 0.1M NaCl, run with 2fs timesteps for a total of 3ns each. We do most of the work at constant pressure and temperature (NPT), using a Langevin piston barostat at $P = 1\text{atm}$, and a Langevin thermostat at $T = 310\text{K}$ and damping rate $\gamma_{\text{L}} = 5\text{ps}^{-1}$. The NPT simulations model the statics well, but the thermostat's damping leads to unphysical dynamics with increased relaxation rates. This allows shorter simulations to equilibrate, but prevents us from determining the rates we should expect to see in reality. We therefore do a second measurement of diffusion at constant volume and energy (NVE).

The center of mass and global rotation of the pair accounts for six trivial degrees of freedom; the remaining six represent the relative position and orientation of the two units. Of these six, only one is a pure translation: the distance $r = |\boldsymbol{r}_2 - \boldsymbol{r}_1|$ between the center of each unit. The orientation of unit $m$ can be represented by a rotation matrix $\boldsymbol{\Omega}_m$ which rotates the unit from its reference orientation by an angle $|\boldsymbol{\theta}_m|$ about the axis $\hat{\boldsymbol{\theta}}_m$. The even and odd combinations $\boldsymbol{\theta}_1 \pm \boldsymbol{\theta}_2$ give six degrees of freedom that comprise the remaining five coordinates, along with an overall rotation due to the even combination about the inter-body axis $\boldsymbol{r}_2 - \boldsymbol{r}_1$.

As we simulate just one pair of units from a protein complex, we omit the forces and torques on them due to the other units in the lattice, which generically had a nonzero resultant, so that the trajectory drifts with time as shown in Figure 6.3(a). In order to expand the free energy around the physiologically relevant configuration, we must add external forces to compensate; in light of (6.1) the correct force to impose is given by $\overline{\boldsymbol{f}} = -\boldsymbol{\Gamma}^{-1}\overline{\boldsymbol{v}}$, where $\overline{\boldsymbol{v}}$ is the drift velocity measured in the absence of the compensating force. Rerunning the simulation from Figure 6.3(a) with the calculated

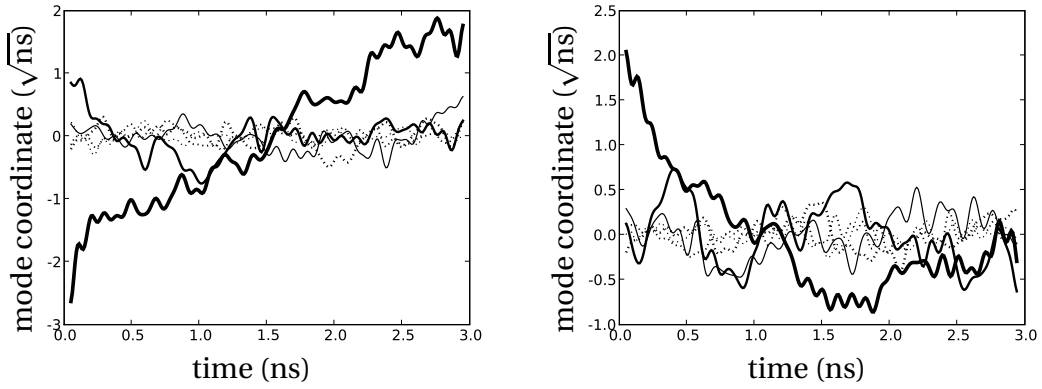



Figure 6.3: (a) Trajectory without a drift-cancelling force. (b) The same simulation with an external force $\boldsymbol{f} = -\boldsymbol{\Gamma}^{-1}\overline{\boldsymbol{v}}$ applied. We see in (a) the slowest mode drifts at a roughly constant rate, and has a computed relaxation time of 1.6ns, while the slowest mode in (b) has a computed relaxation time of 0.5ns, which would be even smaller if we trimmed off (i.e., restarted after) the first ns.

force resulted in the trajectory shown in Figure 6.3(b).

Drift cancellation was not actually important for the pairs reported in our results (the data shown in Figure 6.3 is from a previous simulation that wasn't actually used).

## 6.4 Results

We will start by presenting results at the level of individual simulations, and then go on to aggregate all the simulations to the network level to draw conclusions about the HIV lattice as a whole.

### 6.4.1 Individual simulations

The results for each simulation were similar, and the trajectory of the linker in the transformed relaxation mode coordinates is shown in Figure 6.4, which is characteristic of all the observed trajectories. Once we have an equilibrated segment of a

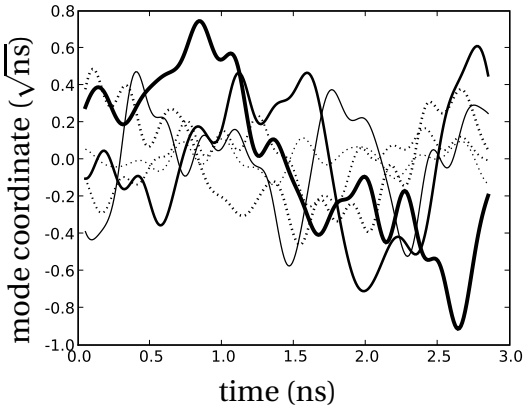



Figure 6.4: Relaxation mode trajectories of linker. The mode coordinate has units of $\sqrt{\text{ns}}$ because it has been normalized by the noise. The slower modes are drawn with thicker lines. Note that the slowest mode has a very small drift, and we could correct this by applying an external force. The traces have been smoothed with a low-pass filter for readability.

trajectory we use (6.4) to determine the $6 \times 6$ stiffness tensor $\mathbf{K}$; different components have different units, so it would be mathematically meaningless to diagonalize it directly. Instead, we define reduced stiffness tensors, representing the free energy cost if we optimize $r$ for a fixed set of angles and vice versa. Given

$$\mathbf{K} = \begin{pmatrix} K_{\mathrm{rr}} & \mathbf{K}_{\mathrm{r}\theta} \\ \mathbf{K}_{\theta\mathrm{r}} & \mathbf{K}_{\theta\theta} \end{pmatrix}, \tag{6.8}$$

then integrating out the orientations (for $K_{\mathrm{stretch}}$) and the stretch (for $\mathbf{K}_{\mathrm{orient}}$) gives reduced tensors

$$K^{(\mathrm{eff})}_{\mathrm{stretch}} = K_{\mathrm{rr}} - \mathbf{K}_{\mathrm{r}\theta}\mathbf{K}^{-1}_{\theta\theta}\mathbf{K}_{\theta\mathrm{r}} \tag{6.9}$$

$$\mathbf{K}^{(\mathrm{eff})}_{\mathrm{orient}} = \mathbf{K}_{\theta\theta} - \mathbf{K}_{\theta\mathrm{r}}K^{-1}_{\mathrm{rr}}\mathbf{K}_{\mathrm{r}\theta}. \tag{6.10}$$

The eigenvalues of these reduced tensors are given in Table 6.1.

We computed the stiffness tensor implicitly in the relative coordinates between the two bodies, but the absolute coordinates are the natural frame for computing the noise. Measuring the diffusion of a single body in an NVE simulation yields a mean

Table 6.1: Effective stiffness eigenvalues for pair simulations: NTD dimer, CTD dimer, and the NTD–CTD linker within the CA protein.

| | $K^{(\mathrm{eff})}_{\mathrm{stretch}}$ ($k_{\mathrm{B}}T/\mathrm{nm}^2$) | $\mathbf{K}^{(\mathrm{eff})}_{\mathrm{orient}}$ eigenvalues ($k_{\mathrm{B}}T$) | | | | |
|---|---|---|---|---|---|---|
| NTD–NTD | 12 | 1300 | 2800 | 4500 | 10000 | 18000 |
| CTD–CTD | 9.9 | 210 | 340 | 1100 | 3900 | 8300 |
| Linker | 2.8 | 130 | 250 | 480 | 1100 | 3800 |

$D^{(\mathrm{rot})}_{\mathrm{CTD}} = 0.11\mathrm{rad}^2/\mathrm{ns}$ and $D^{(\mathrm{rot})}_{\mathrm{NTD}} = 0.044\mathrm{rad}^2/\mathrm{ns}$. If we approximate each domain as a solid sphere then Stokes' law gives a rotational diffusion constant [16]

$$D^{(\mathrm{rot})} = k_{\mathrm{B}}T/(8\pi\eta r^3). \tag{6.11}$$

We thus expect $D^{(\mathrm{rot})}_{\mathrm{CTD}} = 0.11\mathrm{rad}^2/\mathrm{ns}$ and $D^{(\mathrm{rot})}_{\mathrm{NTD}} = 0.050\mathrm{rad}^2/\mathrm{ns}$ using a viscosity $\eta^{(310\mathrm{K})} = 0.69\mathrm{cP}$. The accepted TIP3P viscosity $\eta^{(\mathrm{TIP3P})} = 0.31\mathrm{cP}$ gives poorer agreement.

The translational diffusion constant is slightly harder to measure, since it is influenced significantly by the finite-size effect [17]. This can be corrected for by measuring the diffusion at several box side lengths $L$ and using a linear fit of $D^{(\mathrm{tr})}$ versus $1/L$ to extrapolate to $1/L = 0$. Doing so yields $D^{(\mathrm{tr})}_{\mathrm{CTD}} = 55\text{Å}^2/\mathrm{ns}$ and $D^{(\mathrm{tr})}_{\mathrm{NTD}} = 27\text{Å}^2/\mathrm{ns}$. Stokes' law gives expected $D^{(\mathrm{tr})}_{\mathrm{CTD}} = 56\text{Å}^2/\mathrm{ns}$ and $D^{(\mathrm{tr})}_{\mathrm{NTD}} = 43\text{Å}^2/\mathrm{ns}$ using viscosity $\eta^{(\mathrm{TIP3P})} = 0.31\mathrm{cP}$. The measured $D^{(\mathrm{tr})}$ has a significantly larger relative error than $D^{(\mathrm{rot})}$, due to the finite-$L$ extrapolation.

We can diagonalize the relaxation matrix to compute the relaxation modes for each linkage. The NPT relaxation times from this calculation are listed in Table 6.2. All the times are significantly shorter than the simulation time, so we can be confident that the simulations are equilibrated.

Table 6.2: NPT time constants for the relaxation modes of each pair.

| | relaxation times $\tau_\alpha$ (ps) | | | | | |
|---|---|---|---|---|---|---|
| NTD–NTD | 120 | 23 | 18 | 9.3 | 6.0 | 4.4 |
| CTD–CTD | 76 | 26 | 24 | 7.8 | 5.4 | 4.1 |
| Linker | 190 | 140 | 80 | 76 | 22 | 8.3 |

### 6.4.2 Capsid network

Now that we have characterized the individual springs, can compose them together into a triangular lattice as shown in Figure 6.1, with an NTD hexamer at each vertex, a CTD dimer at the midpoint of each edge, and a spring connecting each domain, whose free energy is given by the relative positions multiplied into the appropriate stiffness tensor. We can then determine the free energy minimum as a function of periodic cell dimensions to find a lattice constant of $a = 9.1$nm. This is slightly smaller than the experimentally measured 10.7nm [8], which may be largely due to our sheet being flat, rather than curved into a tube. Computing the free energy of simple extension yields a two-dimensional Young's modulus of $0.92 k_\mathrm{B}T/\text{Å}^2 = 0.39$N/m and a Poisson ratio of 0.30. Assuming homogeneity and a thickness of 5nm, we find a three-dimensional Young's modulus of 77MPa (compared with 115MPa measured using atomic force microscopy [18]).

Furthermore, we can estimate the relaxation rate of the full-capsid breathing mode in water by further coarse-graining to a single coordinate $a$ representing a uniform dilation in the plane, which has dynamics given by (6.1) with stiffness and mobility constants $K_a$ and $\Gamma_a$. The projected stiffness is given by the bulk modulus $K_a = 4K_{2\mathrm{d}} = 2.6 k_\mathrm{B}T/\text{Å}^2$, calculated from the two-dimensional Young's modulus

and Poisson ratio[2], which we can measure by performing numerical experiments of simple extension and hydrostatic compression using the analytic continuation of the free energy for large deviation from §5.3.6. To simulate an infinite periodic lattice, we use 38 degrees of freedom: $36 = 3 \cdot 12$ to represent the configurations of the three differently oriented proteins (two domains with six degrees of freedom each), and the other two give the periodic cell dimensions $a_x$ and $a_y$. We first find the unconstrained global minimum, and then add stretching constraints to $a_x$ and/or $a_y$, measuring the effect on the new constrained energy minimum.

To project the damping term, we observe that all the actual motion in the breathing mode of a virus capsid of radius $r$ is in the radial direction, and we thus need to scale the capsid protein's translational diffusion constant by $(\mathrm{d}a/\mathrm{d}r)^2$ to find the diffusion constant for $a$. Using the detailed balance condition,

$$\Gamma_a = \frac{16\pi\sqrt{3}}{N} \frac{D_{\mathrm{NTD}}^{(\mathrm{tr})} + D_{\mathrm{CTD}}^{(\mathrm{tr})}}{k_{\mathrm{B}}T} \frac{\eta^{(\mathrm{TIP3P})}}{\eta^{(310\mathrm{K})}}, \tag{6.12}$$

where $N = 16\pi\sqrt{3}r^2/a^2$ is the total number of capsid proteins[3].. Taking $N = 1500$ proteins as the average size for an HIV capsid thus gives a relaxation rate of $6.1\mathrm{ns}^{-1}$ for the breathing mode.

We also attempted to calculate the spontaneous curvature $\theta_0$ and bending stiffness $\kappa$ of a flat sheet of HIV by adding a 39th parameter to the above calculation to represent bending angle for a tube of capsid (rather than a sheet). For the $(m, 0)$ tubes, we found a negative spontaneous curvature, which seems unphysical, and therefore

[2]If we take Lamé coefficients $\lambda$ and $\mu$, defining free energy $F = \frac{1}{2}\lambda u_{ii}^2 + \mu u_{ik}u_{ik}$ then in two dimensions, the bulk modulus $K_{2\mathrm{d}} = \lambda + \mu$, the Young's modulus is $Y_{2\mathrm{d}} = 4K\mu/(K+\mu)$, and the Poisson ratio is $\nu_{2\mathrm{d}} = (K-\mu)/(K+\mu)$. Finally, we can approximate $Y_{3\mathrm{d}} = Y_{2\mathrm{d}}/t$ where $t$ is the thickness.

[3]We have included a term $\eta^{(\mathrm{TIP3P})}/\eta^{(310\mathrm{K})}$ to correct for the TIP3P water model's incorrect viscosity (which affected the diffusion constants we measured in our simulations) so that we can estimate the relaxation rate in real water.

did not proceed to compute the stiffness. Neither did we consider alternate axis directions, since the minimization appeared to be very ill-conditioned.

## 6.5 Discussion

In conclusion, we have reported a model of overdamped random walks in which the statics and dynamics are described respectively by complementary "stiffness" and "mobility" tensors. From these two tensors a "relaxation matrix" can be formed, the eigenvalues of which give the relaxation rates and provide a convergence test for simulations. We demonstrated the usefulness of this model in extracting coarse-grained elastic constants from molecular dynamics trajectories of pairs of interacting units.

While our relaxation formalism is novel in combining stochastic dynamics with a realistic multi-component spring, it does bear some similarities to a number of existing techniques. Normal mode analysis, and in particular, Gaussian network models, replace interactions (either between atoms or groups of atoms) with springs of uniform stiffness [19, 20, 21, 22]. While these techniques have been successful in determining the soft degrees of freedom to explain reaction pathways such as virus maturation [23, 24], the frequencies themselves are well known to be artificial because they omit the damping forces of the surrounding water (this has been addressed by Lamm and Szabo [25] with their "Langevin modes"). Additionally, most of these techniques are insensitive to point mutations or environmental conditions. On the other hand, "essential dynamics" (or "principal component analysis") [26, 27, 28] uses all-atom simulations to determine the soft modes of a system. Moreover, Hayward et al. [29] suggested specifying important modes *a priori*, which is our starting point. We extend these approaches by using the same principal components (to wit, the position and orientation) for each unit to connect the different interactions, and by using the

relaxation modes to predict the dynamics *ab initio*.

Our technique is not specific to virus capsids: a similar approach should be applicable to many other systems of interacting protein domains, such as microtubules or BAR domains. Among virus capsids, HIV was particularly amenable because all the important interactions are pairwise, while many other viruses are complicated by long tails in which all six molecules in the hexamer are entwined. It would be interesting to look at the effect of point mutations, salinity, and pH on the resultant elasticity (which could then be verified by experiment). We hope that the techniques presented here will provide a convenient middle ground between the atomistic and continuum pictures of many biological systems.

## BIBLIOGRAPHY


[1] S. D. Hicks and C. L. Henley, "Coarse-grained protein-protein stiffnesses and dynamics from all-atom simulations", *Phys. Rev. E* **81**, 030903 (2010).

[2] A. Arkhipov, P. L. Freddolino, and K. Schulten, "Stability and Dynamics of Virus Capsids Described by Coarse-Grained Modeling", *Structure* **14**, 1767 (2006).

[3] R. Zandi, D. Reguera, R. F. Bruinsma, W. M. Gelbart, and J. Rudnick, "Origin of icosahedral symmetry in viruses", *Proc. Nat. Acad. Sci. USA* **101**, 15556 (2004).

[4] S. D. Hicks and C. L. Henley, "Irreversible growth model for virus capsid assembly", *Phys. Rev. E* **74**, 031912 (2006).

[5] M. F. Hagan and D. Chandler, "Dynamic Pathways for Viral Capsid Assembly", *Biophys. J.* **91**, 42 (2006).

[6] C. Jarzynski, "Nonequilibrium equality for free energy differences", *Phys. Rev. Lett.* **78**, 2690 (1997).

[7] C. Jarzynski, "Nonequilibrium work relations: foundations and applications", *Eur. Phys. J. B* **64**, 331 (2008).

[8] S. Li, C. P. Hill, W. I. Sundquist, and J. T. Finch, "Image reconstructions of helical assemblies of the HIV-1 CA protein.", *Nature* **407**, 409 (2000).

[9] E. R. Wright, J. B. Schooler, H. J. Ding, C. Kieffer, C. Fillmore, W. I. Sundquist, and G. J. Jensen, "Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells.", *EMBO J.* **26**, 2218 (2007).

[10] B. K. Ganser-Pornillos, A. Cheng, and M. Yeager, "Structure of Full-Length HIV-1 CA: A Model for the Mature Capsid Lattice", *Cell* **131**, 70 (2007).

[11] C. Tang, Y. Ndassa, and M. F. Summers, "Structure of the N-terminal 283-residue fragment of the immature HIV-1 Gag polyprotein", *Nat. Struct. Biol.* **9**, 537 (2002).

[12] G. B. Mortuza, L. F. Haire, A. Stevens, S. J. Smerdon, and J. P. Stoye, "High-resolution structure of a retroviral capsid hexameric amino-terminal domain", *Nature* **431**, 481 (2004).

[13] T. R. Gamble, S. Yoo, F. F. Vajdos, U. K. von Schwedler, D. K. Worthylake, H. Wang, J. P. McCutcheon, W. I. Sundquist, and C. P. Hill, "Structure of the Carboxyl-Terminal Dimerization Domain of the HIV-1 Capsid Protein", *Science* **278**, 849 (1997).

[14] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten, "Scalable molecular dynamics with NAMD.", *J. Comput. Chem.* **26**, 1781 (2005).

[15] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, et al., "All-atom empirical potential for molecular modeling and dynamics studies of proteins", *J. Phys. Chem. B* **102**, 3586 (1998).

[16] H. Lamb, *Hydrodynamics* (Cambridge, 1932), p. 589, 6th ed.

[17] I. C. Yeh and G. Hummer, "Diffusion and electrophoretic mobility of single-stranded RNA from molecular dynamics simulations", *Biophys. J.* **86**, 681 (2004).

[18] N. Kol, Y. Shi, M. Tsvitov, D. Barlam, R. Z. Shneck, M. S. Kay, and I. Rousso, "A Stiffness Switch in Human Immunodeficiency Virus", *Biophys. J.* **92**, 1777 (2007).

[19] M. M. Tirion, "Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis", *Phys. Rev. Lett.* **77**, 1905 (1996).

[20] I. Bahar, A. R. Atilgan, M. C. Demirel, and B. Erman, "Vibrational Dynamics of Folded Proteins: Significance of Slow and Fast Motions in Relation to Function and Stability", *Phys. Rev. Lett.* **80**, 2733 (1998).

[21] F. Tama, M. Valle, J. Frank, and C. L. Brooks, "Dynamic reorganization of the functionally active ribosome explored by normal mode analysis and cryo-electron microscopy", *Proc. Nat. Acad. Sci. USA* **100**, 9319 (2003).

[22] M. M. Gibbons and W. S. Klug, "Mechanical modeling of viral capsids", *J. Mat. Sci.* **42**, 8995 (2007).

[23] A. J. Rader, D. H. Vlad, and I. Bahar, "Maturation Dynamics of Bacteriophage HK97 Capsid", *Structure* **13**, 413 (2005).

[24] E. R. May, (personal communication) has calibrated the stiffnesses in an elastic network model to all-atom MD of the HK97 phage (mature) capsid.

[25] G. Lamm and A. Szabo, "Langevin modes of macromolecules", *J. Chem. Phys.* **85**, 7334 (1986).

[26] T. Horiuchi and N. Go, "Projection of Monte Carlo and molecular dynamics trajectories onto the normal mode axes: human lysozyme", *Proteins* **10**, 106 (1991).

[27] T. Ichiye and M. Karplus, "Collective motions in proteins: A covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations", *Proteins Str. Fun. Gen.* **11**, 205 (1991).

[28] A. Amadei, A. B. Linssen, and H. J. Berendsen, “Essential dynamics of proteins.”, *Proteins* **17**, 412 (1993).

[29] S. Hayward, A. Kitao, and H. J. C. Berendsen, “Model-Free Methods of Analyzing Domain Motions in Proteins From Simulation: A Comparison of Normal Mode Analysis and Molecular Dynamics Simulation of Lysozyme”, *Proteins Str. Fun. Gen.* **27**, 425 (1997).

CHAPTER 7

# DIFFUSION IN MOLECULAR DYNAMICS SIMULATIONS

In our work measuring elastic constants (see Chapter 6), we found it necessary to measure the diffusion constants of our proteins for two reasons: first, in order to scale the stiffness tensor to the more physically meaningful relaxation tensor, we needed to multiply by the diffusion tensor. Secondly, if the relative coordinates are found to drift, the diffusion tensor provides the conversion between the measured drift velocity and the necessary force to cancel the drift.

In attempting to determine the diffusion tensor from the MD trajectories, we learned a number of lessons about diffusion the "hard way". In particular, the effect of box sizes, thermostats, water models, and the interaction between them, while perhaps widely known and passed down in certain circles, are not very well documented in the literature. As a result, we spent a significant amount of time troubleshooting our simulations when the diffusion constants we measured were orders of magnitude different from what we expected to see from simple Stokes drag estimates.

In this chapter we will discuss some preliminary work towards quantifying the corrections due to these factors interacting. If continued, we hope this work will lead to a method for utilizing small simulations with thermostats to efficiently compute diffusion constants.

We begin in §7.1 with a general discussion of stochastic differential equations and how they relate to estimating diffusion constants from simulated trajectories.

After laying this groundwork, we explore in §7.2 the primary factors that cause the the measured diffusion constants to differ from their actual physical values. We outline a plan for quantifying the effect of these factors in §7.3, and present some early results in §7.4.

## 7.1 Measuring diffusion

Here we fill in the computational details that we left out of §6.2 for measuring the diffusion tensor. In §7.1.1 we write the diffusion tensor as a linear fit $\mathbf{W}(t) = 2\mathbf{D}t$, where $\mathbf{W}(t)$ is given by time averages of the trajectory. We then show that the uncertainty in the measurement of $\mathbf{W}(t)$ scales as $t^{3/2}$, suggesting we should weight the points differently in the linear fit to compute $\mathbf{D}$. In §7.1.2 we write $\mathbf{W}(t)$ in terms of Fourier transforms, adding a correction to the Wiener-Khintcine theorem to deal properly with the case of finite-time trajectories. Finally, we take an aside in §7.1.4 to discuss the optimal way to measure the drift velocity of a trajectory.

### 7.1.1 The diffusion tensor

We were not particularly careful in §6.2 to specify which degrees of freedom we were measuring the diffusion of. While it's possible to calculate a diffusion tensor for the relative degrees of freedom, it's not particularly meaningful. Thus, we concentrate in this section on the absolute coordinates of a single body moving freely in space, such that the velocity

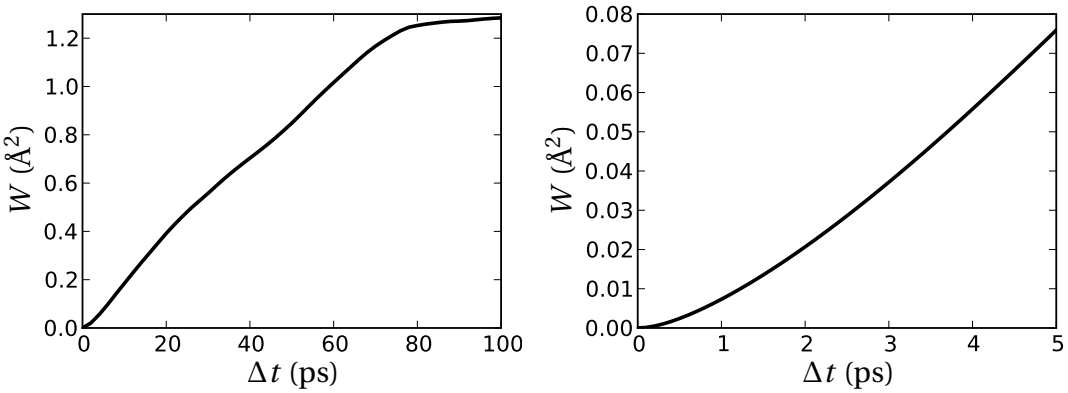
$$\boldsymbol{v}(t) = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} \tag{7.1}$$

is white noise,

$$\left\langle \boldsymbol{v}(t)\boldsymbol{v}(t')\right\rangle_{\boldsymbol{\zeta}} = 2\mathbf{D}\delta(t-t'). \tag{7.2}$$

We've refined the notation $\langle\cdot\rangle_{\boldsymbol{\zeta}}$ to mean an ensemble average, as contrasted with $\langle\cdot\rangle_t$ which will signify a time average $\langle f\rangle_t = \frac{1}{t_{\mathrm{run}}}\int_0^{t_{\mathrm{run}}}\mathrm{d}t\, f(t)$, where $t_{\mathrm{run}}$ is the total run time.

This approximation of white noise is good for time scales longer than the ballistic scale ($1/\gamma$, as appearing in (7.23)), as seen indirectly in Figure 7.1(b). Additionally it's




(a) (b)

Figure 7.1: Plot of the autocorrelation $W(\Delta t) = \mathrm{Tr}\mathbf{W}(\Delta t)$ at two different time scales. (a) 100ps, and (b) 5ps. Note that on scales from 5ps to about 50ps the curve is roughly linear, while we can see the (quadratic) ballistic behavior at scales shorter than 1 or 2ps.

only reasonable for infinite fluids, for reasons we'll describe in §7.2.

We gave a recipe in (6.5) to measure $\mathbf{D}$ by calculating the autocorrelation $\mathbf{W}(\Delta t)$ using a time average of the trajectory:

$$\mathbf{W}(\Delta t) \equiv \Big\langle \big(\boldsymbol{r}(t+\Delta t) - \boldsymbol{r}(t)\big)\big(\boldsymbol{r}(t+\Delta t) - \boldsymbol{r}(t)\big) \Big\rangle_t \to 2\mathbf{D}\Delta t \tag{7.3}$$

Indeed, expanding this time average gives

$$\big\langle \mathbf{W}(\Delta t) \big\rangle_{\boldsymbol{\zeta}} = \frac{1}{t_{\mathrm{run}}} \int_0^{t_{\mathrm{run}}} \mathrm{d}t \int_t^{t+\Delta t} \mathrm{d}\tau \int_t^{t+\Delta t} \mathrm{d}\tau' \big\langle \boldsymbol{v}(\tau)\boldsymbol{v}(\tau') \big\rangle_{\boldsymbol{\zeta}} = 2\mathbf{D}\Delta t, \tag{7.4}$$

as expected. An example of $W(\Delta t) = \mathrm{Tr}\mathbf{W}(\Delta t)$ is plotted in Figure 7.1.

We can go further, however, and determine the uncertainty in our calculation of $\mathbf{W}(\Delta t)$ by computing

$$\begin{aligned} \big\langle \mathbf{W}_{ij}(\Delta t)^2 \big\rangle_{\boldsymbol{\zeta}} = \frac{1}{{t_{\mathrm{run}}}^2} \int_0^{t_{\mathrm{run}}} \mathrm{d}t_1\, \mathrm{d}t_2 \int_{t_1}^{t_1+\Delta t} \mathrm{d}\tau_1\, \mathrm{d}\tau_1' \\ \int_{t_2}^{t_2+\Delta t} \mathrm{d}\tau_2\, \mathrm{d}\tau_2' \big\langle \boldsymbol{v}_i(\tau_1)\boldsymbol{v}_j(\tau_1')\boldsymbol{v}_i(\tau_2)\boldsymbol{v}_j(\tau_2') \big\rangle_{\boldsymbol{\zeta}}. \end{aligned} \tag{7.5}$$

(Note that we are *not* using Einstein summation notation here: $i$ and $j$ are free indices throughout). Wick's theorem allows separating the four-way expectation into three

pairs of expectations of pairs:

$$\begin{aligned}\left\langle \boldsymbol{v}_i(\tau_1)\boldsymbol{v}_j(\tau_1')\boldsymbol{v}_i(\tau_2)\boldsymbol{v}_j(\tau_2')\right\rangle &= \left\langle \boldsymbol{v}_i(\tau_1)\boldsymbol{v}_j(\tau_1')\right\rangle\left\langle \boldsymbol{v}_i(\tau_2)\boldsymbol{v}_j(\tau_2')\right\rangle \\ &+ \left\langle \boldsymbol{v}_i(\tau_1)\boldsymbol{v}_j(\tau_2')\right\rangle\left\langle \boldsymbol{v}_j(\tau_1')\boldsymbol{v}_i(\tau_2)\right\rangle + \left\langle \boldsymbol{v}_i(\tau_1)\boldsymbol{v}_i(\tau_2)\right\rangle\left\langle \boldsymbol{v}_j(\tau_1')\boldsymbol{v}_j(\tau_2')\right\rangle.\end{aligned} \tag{7.6}$$

The first term separates into uncoupled integrals leaving simply $(2\mathbf{D}_{ij}\Delta t)^2$. The second and third terms are slightly more complicated. We rearrange the integrals in (7.5) to compute $\sigma^2_{\mathbf{W}_{ij}(\Delta t)} \equiv \left\langle \mathbf{W}_{ij}(\Delta t)^2\right\rangle_{\boldsymbol{\zeta}} - \left\langle \mathbf{W}_{ij}(\Delta t)\right\rangle^2_{\boldsymbol{\zeta}}$:

$$\sigma^2_{\mathbf{W}_{ij}(\Delta t)} = 4\frac{\mathbf{D}_{ii}\mathbf{D}_{jj}+\mathbf{D}_{ij}^2}{{t_{\text{run}}}^2}\int_0^{t_{\text{run}}} \mathrm{d}t_1\,\mathrm{d}t_2\,F(t_1,t_2)^2, \tag{7.7}$$

where

$$F(t_1,t_2) = \int_{t_1}^{t_1+\Delta t}\mathrm{d}\tau_1\int_{t_2}^{t_2+\Delta t}\mathrm{d}\tau_2\delta(\tau_1-\tau_2). \tag{7.8}$$

We can evaluate $F$ piecewise to find[1]

$$F(t_1,t_2) =_{\max}\left(0,\Delta t - |t_1-t_2|\right). \tag{7.9}$$

We can change variables in (7.7) by letting $u = \frac{1}{2}(t_1+t_2)$ and $v = t_1 - t_2$. Now $u$ goes from 0 to $t_{\text{run}}$ and the integration range of $v$ is a function of $u$. But $F(v)$ vanishes for large $|v|$ so we can take $v = -\infty\ldots\infty$ without loss. Thus,

$$\begin{aligned}\sigma^2_{\mathbf{W}_{ij}(\Delta t)} &= 4\frac{\mathbf{D}_{ii}\mathbf{D}_{jj}+\mathbf{D}_{ij}^2}{t_{\text{run}}}\int_{-\infty}^{\infty}\mathrm{d}v\,F(v)^2 \\ &= \frac{8\Delta t^3}{3t_{\text{run}}}\left(\mathbf{D}_{ii}\mathbf{D}_{ij}+\mathbf{D}_{ij}^2\right).\end{aligned} \tag{7.10}$$

If we take $\mathbf{D}_{ij} = D\delta_{ij}$ then

$$\sigma^2_{\mathbf{W}(\Delta t)} = \frac{16D^2\Delta t^3}{3t_{\text{run}}}. \tag{7.11}$$

[1]Note that we could have averaged $\mathbf{W}(\Delta t_1)\mathbf{W}(\Delta t_2)$ (taking $\Delta t_1 < \Delta t_2$ without loss of generality) and the only difference would be a plateau of height $\Delta t_1$ and width $\Delta t_2 - \Delta t_1$ in $F(t_1 - t_2)$ such that

$$\int_{-\infty}^{\infty}\mathrm{d}vF(v)^2 = \Delta t_1^2\left(\Delta t_2 - \frac{1}{3}\Delta t_1\right).$$

So we see that our measurement $\mathbf{W}(\Delta t)$ has an uncertainty scaling with $\Delta t^{3/2}$ and we can therefore weight our linear fit accordingly.

### 7.1.2 Fourier transforms

While we can compute $\mathbf{W}(\Delta t)$ separately for each value of $\Delta t$ by computing the average $\left(\boldsymbol{r}(t+\Delta t)-\boldsymbol{r}(t)\right)^2$ for all $t = 0 \ldots t_{\text{run}} - \Delta t$, this calculation scales quadratically with runtime, which quickly becomes impractical (particularly if we take a timestep of 2fs). Instead we take advantage of the fast Fourier transform, which scales as $O(N \log N)$, a big improvement.

If our trajectories were periodic (or infinite), the Wiener-Khintcine theorem makes this computation straightforward. We compute the power spectrum

$$\widetilde{\mathbf{S}}_{ij}(\omega) = \frac{1}{t_{\text{run}}} \widetilde{\boldsymbol{r}}_i(\omega) \widetilde{\boldsymbol{r}}_j^*(\omega) \tag{7.12}$$

and then compute the inverse transform to find the autocorrelation

$$\mathbf{W}(\Delta t) = \mathbf{S}(\Delta t). \tag{7.13}$$

In our case, however, we need to be more careful, since we have neither periodic nor infinite boundary conditions, but instead compute

$$\begin{aligned} \mathbf{W}_{ij}(\Delta t) = \frac{1}{t_{\text{run}} - \Delta t} \int_0^{t_{\text{run}}-\Delta t} \mathrm{d}t \Big( & \boldsymbol{r}_i(t+\Delta t)\boldsymbol{r}_j(t+\Delta t) + \boldsymbol{r}_i(t)\boldsymbol{r}_j(t) \\ & - \boldsymbol{r}_i(t+\Delta t)\boldsymbol{r}_j(t) - \boldsymbol{r}_i(t+\Delta t)\boldsymbol{r}_j(t) \Big). \end{aligned} \tag{7.14}$$

Notice that if we define boundary conditions $\boldsymbol{r}(t) = 0$ for $t < 0$ and $t > T$ then the Fourier transforms from $-\infty$ to $\infty$ are still well-defined, and we find

$$\mathbf{S}_{ij}(\Delta t) = \int_0^{t_{\text{run}}-\Delta t} \mathrm{d}t\, \boldsymbol{r}_i(t+\Delta t)\boldsymbol{r}_j(t). \tag{7.15}$$

Thus,

$$\begin{aligned}\mathbf{W}_{ij}(\Delta t) = \frac{1}{t_{\text{run}} - \Delta t}\Big(2\mathbf{S}_{ij}(0) - \mathbf{S}_{ij}(\Delta t) - \mathbf{S}_{ji}(\Delta t) \\ - \int_0^{\Delta t} \mathrm{d}\tau \Big(\boldsymbol{r}_i(\tau)\boldsymbol{r}_j(\tau) + \boldsymbol{r}_i(t_{\text{run}} - \tau)\boldsymbol{r}_j(t_{\text{run}} - \tau)\Big)\Big),\end{aligned} \tag{7.16}$$

where the subtracted integral is to correct for the fact that $\boldsymbol{r}(t+\Delta t)\boldsymbol{r}(t+\Delta t) + \boldsymbol{r}(t)\boldsymbol{r}(t)$ counts points close to 0 and $t_{\text{run}}$ once each, but counts points in the middle twice each. This gives an efficient algorithm for computing the diffusion tensor from a trajectory. Note that $\mathbf{S}_{ij}(-\Delta t) = \mathbf{S}_{ji}(\Delta t)$, so that $\mathbf{S}(\Delta t)$ is nonzero for a length $2t_{\text{run}}$. We therefore need to pad $\boldsymbol{r}(t)$ with at least as many zeros as we have actual data points.

### 7.1.3 Orientational diffusion

While calculating the power spectrum is simple enough with translational diffusion, it becomes more complicated if we're dealing with orientations, since we cannot simply subtract $\boldsymbol{\Theta}(t+\Delta t) - \boldsymbol{\Theta}(t)$ as we could with translations. Rather, we define

$$\boldsymbol{\Delta\Theta}(t, t+\Delta t) = \boldsymbol{\Theta}(t)^{\mathsf{T}}\boldsymbol{\Theta}(t+\Delta t). \tag{7.17}$$

If we represent $\boldsymbol{\Theta}(t)$ as a quaternion $\mathfrak{q}(t) = \rho(t) - \mathrm{i}\boldsymbol{\zeta}(t)\cdot\sigma$ then we write

$$\begin{aligned}\boldsymbol{\Delta\Theta}(t,t+\Delta t) \to \big(\rho(t)\rho(t+\Delta t) + \boldsymbol{\zeta}(t)\boldsymbol{\zeta}(t+\Delta t)\big) \\ - \mathrm{i}\sigma\cdot\big(\rho(t)\boldsymbol{\zeta}(t+\Delta t) - \rho(t+\Delta t)\boldsymbol{\zeta}(t) - \boldsymbol{\zeta}(t)\times\boldsymbol{\zeta}(t+\Delta t)\big).\end{aligned} \tag{7.18}$$

Multiplying (the vector part) by $\boldsymbol{r}(t+\Delta t) - \boldsymbol{r}(t)$ results in ten terms, but each term separates into, e.g. $\big(\boldsymbol{r}(t+\Delta t)\rho(t+\Delta t)\big)\boldsymbol{\zeta}(t)$, which is still a correlation function. Likewise for the θ–θ autocorrelation.

Finally, one other concern when dealing with rotations is that we would like to ultimately represent the translational diffusion tensor in relative coordinates to the object's orientation, rather than in the global frame. If the object rotates significantly during the simulation, we would need a more clever technique to ensure

that all the averaged vectors were in comparable frames. Alternately, if the rotational correlation time is short compared to the simulation time, we end up with an isotropic (rotationally-averaged) diffusion tensor. We never attempted to measure any anisotropy in the diffusion tensor.

### 7.1.4 The drift velocity

If we suppose the equation of motion has a constant drift velocity term,

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \boldsymbol{v} + \boldsymbol{\zeta}(t), \tag{7.19}$$

then it is obvious that $\langle \boldsymbol{r}(t) \rangle_{\boldsymbol{\zeta}} = \boldsymbol{v}t$. But what is the most accurate way to determine $\boldsymbol{v}$ from a given trajectory $\boldsymbol{r}(t)$? For convenience let us take $t = -t_{\mathrm{run}} \ldots t_{\mathrm{run}}$ and define $\boldsymbol{v}_{\mathrm{est}}(t) = \frac{\boldsymbol{r}(t)-\boldsymbol{r}(-t)}{2t}$ to be a reasonable estimate of the velocity from two points. It is easy to see that $\langle \boldsymbol{v}_{\mathrm{est}}(t) \rangle_{\boldsymbol{\zeta}} = \boldsymbol{v}$ for all $t$. There are two very reasonable ways to aggregate the data from the entire trajectory:

$$\boldsymbol{v}_{\mathrm{fit}} \equiv \frac{\int_{-t_{\mathrm{run}}}^{t_{\mathrm{run}}} \boldsymbol{v}_{\mathrm{est}}(t)\, t^2 \,\mathrm{d}t}{\int_{-t_{\mathrm{run}}}^{t_{\mathrm{run}}} t^2 \,\mathrm{d}t}, \tag{7.20}$$

or else

$$\boldsymbol{v}_{\mathrm{end}} \equiv \boldsymbol{v}_{\mathrm{est}}(t_{\mathrm{run}}). \tag{7.21}$$

We can perform a similar Bayesian analysis as in the previous section §7.1.1, with the lemma $\langle \boldsymbol{v}_{\mathrm{est}}(t)\, \boldsymbol{v}_{\mathrm{est}}(t') \rangle_{\boldsymbol{\zeta}} = \frac{\mathbf{D}\min(t,t')}{tt'}$ to see that

$$\langle \boldsymbol{v}_{\mathrm{fit}}^2 \rangle_{\boldsymbol{\zeta}} - \langle \boldsymbol{v}_{\mathrm{fit}} \rangle_{\boldsymbol{\zeta}}^2 = \frac{6D}{5t_{\mathrm{run}}}, \tag{7.22}$$

while $\langle \boldsymbol{v}_{\mathrm{end}}^2 \rangle_{\boldsymbol{\zeta}} - \langle \boldsymbol{v}_{\mathrm{end}} \rangle_{\boldsymbol{\zeta}}^2 = D/t_{\mathrm{run}}$. Surprisingly, we get a better estimate of the drift velocity using just the end points.

## 7.2 Factors affecting diffusion measurements

Here we look at each factor as it individually affects measurements of diffusion in molecular dynamics simulations.

### 7.2.1 Water model

While molecular dynamics has gained a lot of traction for its usefulness in measuring static quantities, it is far less trusted in accurately measuring dynamics. In particular, the two predominant water models, TIP3P (a three-point model) and SPC/E (a simple point charge model), are both widely acknowledged to do a terrible job in simulating something as simple as the self-diffusion of water. Mark and Nilsson [1] measured $D = 0.59 \times 10^{-8} \mathrm{m}^2/\mathrm{s}$ for TIP3P water, contrasted with the experimentally measured $D = 0.23 \times 10^{-8} \mathrm{m}^2/\mathrm{s}$. There are a number of specializations of these water models, but the cost of correcting the dynamics is perhaps paid in less accurate protein-water interactions. Alternately, models such as TIP4P (which extends TIP3P by adding an extra point to keep track of the oxygen's lone pairs; there is also a TIP5P) are able to improve the dynamics instead at the cost of a significantly higher computational burden.

The methods we described in Chapter 6 make use of diffusion measurements, and the relaxation times that come from them, in two ways: first, to determine whether a given simulation has a long enough runtime to assure it has been fully equilibrated; second, to make physical claims about the relaxation time of the actual system in water. The choice of water model is irrelevant to the first consideration, since we are concerned only in how long it takes the simulation to equilibrate, and not how long it would take the physical system to do so. The second consideration, on the other

hand, requires not only an accurate water model, but would also be affected by the differences between pure water and cytosol (at least for biomolecules *in vivo*), and so ad-hoc corrections made to the diffusion after the fact are about as good as we can hope to achieve in these situations anyway.

### 7.2.2 Thermostats

Two other factors are much easier to control: the thermostat and the size of the periodic cell.

It is common practice in molecular dynamics simulations of biological systems to run in a canonical ensemble, holding temperature fixed (typically) somewhere between 297K and 310K. There are a number of different thermostats used for this: Langevin, Berendsen, and Nosé–Hoover. However, they all work by uniformly adding energy into the system in the form of a stochastic force, and subtracting energy from the system in the form of a velocity-dependent damping force. This is in fact the same effect that the water has on a protein that causes its dynamics (at times on the order of picoseconds) to be well-described by a random walk. Thus, if the thermostat is forcing and damping the system in addition to the water, we should expect to measure different dynamics than we would in a microcanonical ensemble with no thermostat. With a sufficiently weak thermostat, we should expect to recover the microcanonical result.

Indeed, Dünweg [2] shows mathematically that small couplings recover the microcanonical ensemble, yet concludes that the screening caused by the thermostat renders simulations in the canonical ensemble useless (because in order for the screening to be sufficiently weak, one must turn the thermostat down far enough that the simulation might as well have been microcanonical to begin with). On the

other hand, since we are not directly concerned with preserving the microcanonical details, but rather inferring them with as little computational power as possible, we shall proceed.

Because the thermostat is effected by a pair of stochastic and frictional forces, we must consider a second order equation of motion in order to reason about it. We take

$$m\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = -m\boldsymbol{\gamma}\boldsymbol{v} + \boldsymbol{\xi}(t), \tag{7.23}$$

where $\boldsymbol{\gamma} = \boldsymbol{\gamma}_\mathrm{h} + \boldsymbol{\gamma}_\mathrm{t}$ is the friction, caused by a combination of the hydrodynamics and the thermostat, and $\boldsymbol{\xi}(t)$ is a stochastic noise function satisfying

$$\left\langle \boldsymbol{\xi}(t)\boldsymbol{\xi}(t')\right\rangle = 2\boldsymbol{\gamma} m k_\mathrm{B} T\delta(t-t') \tag{7.24}$$

by detailed balance. Multiplying by $\frac{1}{m}\mathrm{e}^{\boldsymbol{\gamma}t}$ allows us to integrate, yielding

$$\boldsymbol{v}(t) = \frac{1}{m}\mathrm{e}^{-\boldsymbol{\gamma}t}\int_{-\infty}^{t}\boldsymbol{\xi}(t')\mathrm{e}^{\boldsymbol{\gamma}t'}\,\mathrm{d}t', \tag{7.25}$$

so that

$$\left\langle \boldsymbol{v}(t)\boldsymbol{v}(t')\right\rangle = \frac{k_\mathrm{B}T}{m}\mathrm{e}^{-\boldsymbol{\gamma}|t-t'|}. \tag{7.26}$$

We can relate this to $\mathbf{D}$ by the Einstein relation[2],

$$2\mathbf{D} = \int_{-\infty}^{\infty}\mathrm{d}t\left\langle \boldsymbol{v}(0)\boldsymbol{v}(t)\right\rangle = 2\frac{k_\mathrm{B}T}{m}\left(\boldsymbol{\gamma}_\mathrm{h}+\boldsymbol{\gamma}_\mathrm{t}\right)^{-1}. \tag{7.27}$$

So we see that the diffusion constants from linearly combined effects add reciprocally. More generally, we might conjecture that the effect of adding a thermostat $\gamma$ is

[2]To see this, note that $\boldsymbol{x}(t)-\boldsymbol{x}(0) = \int_0^t \mathrm{d}t'\,\boldsymbol{v}(t')$, so that as $t\to\infty$,

$$\begin{aligned} 2\mathbf{D}t = \left\langle \left(\boldsymbol{x}(t)-\boldsymbol{x}(0)\right)^2\right\rangle &= \int_0^t \mathrm{d}t'\int_0^t \mathrm{d}t''\left\langle \boldsymbol{v}(t)\boldsymbol{v}(t)\right\rangle \\ &\to \int_0^t \mathrm{d}t'\int_{-\infty}^{\infty}\mathrm{d}\tau\left\langle \boldsymbol{v}(t')\boldsymbol{v}(t'+\tau)\right\rangle = t\int_{-\infty}^{\infty}\mathrm{d}\tau\left\langle \boldsymbol{v}(0)\boldsymbol{v}(\tau)\right\rangle \end{aligned}$$

since $\left\langle \boldsymbol{v}(t')\boldsymbol{v}(t'')\right\rangle \to 0$ for large $|t'-t''|$.

to multiply $\langle \boldsymbol{v}(t)\boldsymbol{v}(0)\rangle \to \langle \boldsymbol{v}(t)\boldsymbol{v}(0)\rangle e^{-\gamma t}$. We can therefore associate a time scale of $1/\gamma$ with the thermostat (ranging from 0.1 to $10^3$ps in our simulations). A common value for general MD simulations is $\gamma = 5/\text{ps}$.

### 7.2.3 Finite size effect

The periodic cell size is well-known to affect diffusion by a hydrodynamic finite-size effect [3]. The widely-accepted correction, $D^{(\text{fse})} = D^{\infty} - C/L$, where $C$ is some constant and $L$ is the average box length, is used rather often [4, 5, 6, 7], but clearly it must break down for small enough $L$ that $D^{(\text{fse})}$ would be negative. Because our preference in simulations is to stretch the limits of economy, we have in fact come close to this breakdown.

The physics behind the finite size effect is somewhat subtle. First, consider the effect of an infinite water bath on a protein. We can model this effect as a stochastic force $\boldsymbol{\xi}(t)$ together with a damping force $-\boldsymbol{\gamma}\boldsymbol{v}(t)$ proportional to the protein's velocity. The protein therefore reacts back on the water with a force in the same direction as its velocity. We can then think of this impulse traveling on a random walk and potentially returning to its point of origin, the protein, delivering back the impulse.

In an infinite three-dimensional water bath, the probability of a random walk returning to the origin is rather small. On the other hand, the probability of a random walk returning to any *periodic image* of the origin is much greater, and we can expect it to occur sooner when it does. Thus, the response function is more pronounced. In practice, this has the net effect of decreasing $D$, although we haven't yet worked out the details behind this.

The time scale for the finite size effect is given by the kinematic viscosity of water, $\nu = \eta/\rho \approx 10^{-6}\text{m}^2/\text{s}$. For a box with length $L$ we can expect the finite-size features of

the long-time tail to have a time scale of $L^2/\nu$, which ranges from 4 to 25ps for our box sizes.

### 7.2.4 Finite run time

One final effect to consider is that simulations must have a finite run time $t_{\text{run}}$. This run time necessarily imposes a cut-off in the integral

$$\mathbf{D} = \int_0^{t_{\text{run}}} \left\langle \boldsymbol{v}(t)\boldsymbol{v}(0) \right\rangle \mathrm{d}t \tag{7.28}$$

so that long-time effects are not observed. Moreover, the upper limit is effectively much shorter than $t_{\text{run}}$ since the quality of the correlation data (cf. (7.11)) decays rapidly at the extremes where there are very few pairs of measurements $0 < t < t + \delta t < t_{\text{run}}$ to average over when $\delta t \lesssim t_{\text{run}}$.

## 7.3 Methods

In this section we describe some simulations we performed to measure quantitatively the effect of both the thermostat and the box size (particularly in the small size limit) on the measured diffusion constant.

We performed two different preliminary diffusion simulations to investigate the interactions of the effects discussed in §7.2.

We started with simulations of a single polypeptide in a periodic water box. We chose the smallest polypeptide that could possibly be considered a protein—chignolin—consisting of only ten amino acids. Even at such a small size, the necessary computational resources were significant, and it became clear that we would be unable to do the kind of exploratory work we needed with such slow results.

Our later simulations consisted only of the water box, measuring the self-diffusion of the water molecules. The advantage of the self-diffusion simulations is that in a box of $N$ molecules, we are effectively performing $N$ simultaneous diffusion simulations, since each water molecule is diffusing separately (in practice the effective efficiency boost is smaller, since neighboring molecules are correlated). One important benefit of this is that the accuracy of the results is largely independent of the size of the box, if the different-sized boxes are simulated for the same CPU time, since the time it takes to compute each frame scales with the number of molecules. So holding CPU time constant, a larger box will have fewer frames, but more molecules to average over.

We chose ten different cubical box sizes, with side lengths $L = 20\text{Å}\ldots50\text{Å}$, roughly evenly spaced in $1/L$. We filled each box with TIP3P water and ran a short simulation at constant temperature (297K) and constant pressure (1atm) to calibrate the appropriate box size for the fixed-volume simulations[3]. We then run six simulations at each of the ten fixed volumes: one with no thermostat (NVE, or constant energy), and the other five at 297K (NVT) with thermostat strengths varying from 0.001/ps to 10/ps. The run time varied from 50ps in the largest box to 850ps in the smallest box.

## 7.4 Results and discussion

We first look at the (translational) diffusion results from all sixty simulations. Afterwards, we'll discuss some of the factors that influenced and hindered these results.

[3]Note that fixed-pressure simulations with periodic boundaries cannot provide accurate diffusion measurements because the changing boundary conditions make it impossible to compare distances at different times.

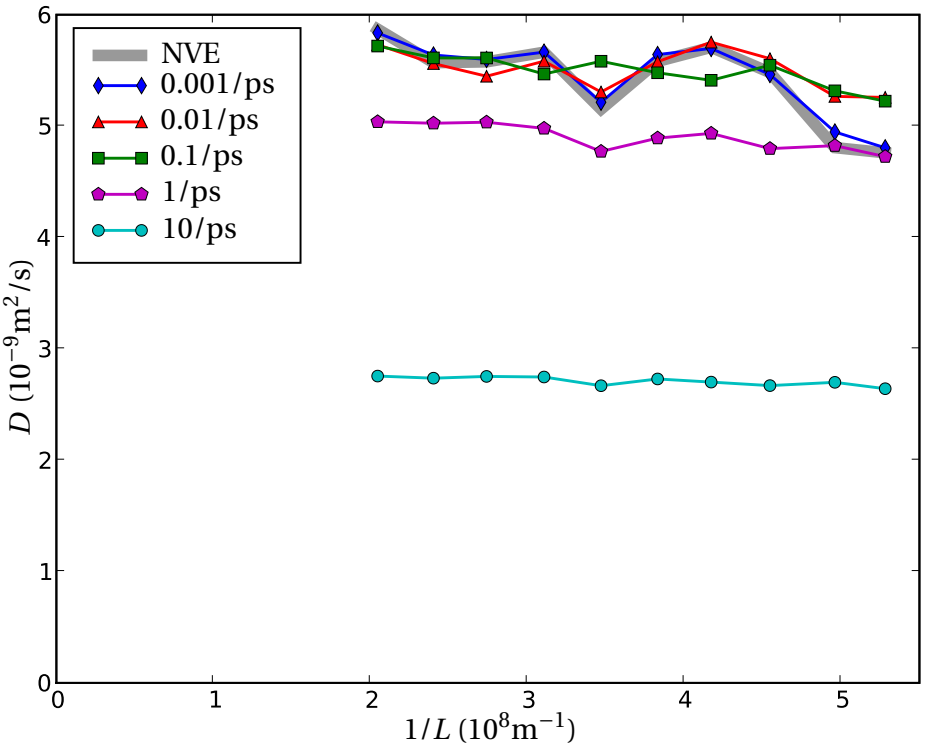



Figure 7.2: Self-diffusion of a periodic box of TIP3P water as a function of box volume $L^3$ and thermostat damping rate parameter $\gamma$ (in units of 1/ps). An accepted value for $D$ for TIP3P water is $5.9 \times 10^{-9}$m$^2$/s [1].

### 7.4.1 Diffusion constants

We plot the raw diffusion constant results in Figure 7.2, and can see a number of interesting features. First we note that, as expected, the 0.001/ps data were nearly identical to the NVE data, and the 0.01/ps were close behind. Of course as the thermostat strength goes to zero, the NVT data should approach NVE. But we do see a significant decrease in the diffusion constant due to damping at the largest thermostat strengths ($\gamma \geq 1$/ps), in accord with (7.27).

Perhaps the most interesting thing to note is the right-hand part of the 0.1/ps line, which has a diffusion constant greater than the NVE line at that volume. In-

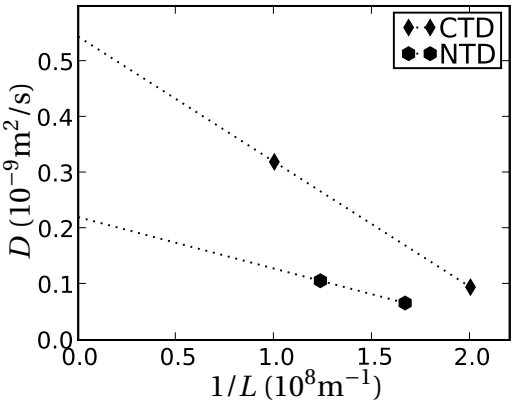



Figure 7.3: Diffusion constants measured from NVE simulations of a single diffusing HIV CTD (mass 8kDa) and NTD (mass 16kDa) at two different box sizes each, as used to calculate the translational diffusion constants in §6.4.1. The finite size effect is much more pronounced here than in Figure 7.2.

deed, the finite-size effect appears to be less pronounced in simulations with stronger thermostats. Moreover, the finite size effect is also generally less pronounced in self-diffusing water than in the diffusion of larger objects, such as proteins, as seen in Figure 7.3.

### 7.4.2 Run time considerations

One problem with Figure 7.2 is that we did not keep the simulation time constant across different points along the $x$-axis, making this plot difficult to interpret properly. Moreover, we can construct a unitless constant $\gamma t_{\text{run}}$, which is different at every single point, varying from $10^{-2}$ at the top-left, to $10^4$ at the bottom-right.

Indeed, the two distinctly separate curves ($\gamma$ = 1/ps and 10/ps) both have $\gamma t_{\text{run}} \gg 1$, while the lines that are indistinguishable from NVE ($\gamma = 0.001$/ps, as well as the left-hand part of $\gamma = 0.01$/ps) have $\gamma t_{\text{run}} \ll 1$. This is reasonable, since $\gamma t_{\text{run}} \ll 1$ means that the thermostat has not yet had time to act. Future studies should take $t_{\text{run}}$ into consideration when choosing values for $\gamma$, so that all simulations run for times

longer than $1/\gamma$.

### 7.4.3 Correlations

Another feature of the data in Figure 7.2 is that for a given box size and across all time, there are correlations in the diffusion constant (that is, they all increase and decrease together). See, in particular, the nearly identical decrease in $D$ for all values of $\gamma \leq 0.01/\text{ps}$ at $1/L \approx 3.5 \times 10^8/\text{m}$. We should have expected this, since the only source of randomness is the Langevin thermostat (the simulated hydrodynamic damping is deterministic), and at small $\gamma$, the thermostat has almost no effect. In future simulations, we should randomize the velocities and pre-equilibrate each simulation separately, causing such correlations to disappear.

We can also get an idea of the uncertainty in these short measurements. A longer run time would certainly be required for production-quality data.

### 7.4.4 Center of mass drift

If we don't renormalize the coordinates of the water molecules to always choose the image inside the primary periodic cell and then replay the trajectories, we find a curious phenomenon: the center of mass of the system moves over time. Indeed, we expect that the non-momentum-conserving Langevin thermostat will cause center of mass diffusion, with diffusion constant scaled inversely by the number of molecules:

$$D_\gamma = \frac{k_\text{B} T}{\rho \gamma_\text{t} V}, \tag{7.29}$$

where $\rho$ is the density of water and $V$ is the volume of the box. We can then estimate the drift velocity variance after a time $t_\text{run}$

$$\left\langle v_\text{d}^2 \right\rangle = \frac{D_\gamma}{t_\text{run}}. \tag{7.30}$$

Table 7.1: Center of mass average drift velocity $v_d$ from one simulation, as a function of inverse box size $1/L$ and thermostat strength $\gamma$.

| | $5.3\times10^8$/m | $4.2\times10^8$/m | $3.1\times10^8$/m | $2.0\times10^8$/m |
|---|---|---|---|---|
| NVE | 1.8m/s | 0.67m/s | 0.48m/s | 0.14m/s |
| 0.001/ps | 20m/s | 13m/s | 2.6m/s | 2.2m/s |
| 0.1/ps | 6.7m/s | 8.3m/s | 11m/s | 5.2m/s |
| 10/ps | 0.69m/s | 0.79m/s | 0.51m/s | 0.57m/s |

Since our simulations kept $V t_{\text{run}} \approx 6.5\text{nm}^3\text{ns}$ roughly constant, we should expect an rms drift velocity of

$$v_{\text{rms}} \approx 1.6(\gamma \cdot \text{ps})^{-1/2}\text{m/s}. \tag{7.31}$$

The average drift velocity for some of the simulations is shown in Table 7.1. We can see that at large $\gamma$ the drift is roughly what we predicted in (7.31). What is surprising, though, is the difference between $\gamma = 0.001/\text{ps}$ and the NVE simulations. The drift at 0.001/ps is smaller than we'd expect, and the wide differences between different volumes is presumably due to the transition from $\gamma t_{\text{run}} \sim 1$ to $\gamma t_{\text{run}} \ll 1$. On the other hand, in the NVE ensemble, if we had started with zero net momentum we would expect no drift at all. Since we did not zero the momentum at the beginning, we expect a ballistic trajectory, but this is surprisingly not what we observed. Rather, the center of mass drifted along a very erratic path. The necessary accelerations are most likely due to small numerical errors in the partial mesh Ewald summation used to efficiently calculate the electrostatics (and these would be present even if we used a momentum-conserving thermostat [7, 8, 9]).

One unfortunate consequence of this drift is that it distorts the diffusion calculation that we cannot easily counteract without repeating the simulations. In these

repeated simulations we can cancel this drift by first starting with an initial condition with no net momentum and then periodically removing any center of mass drift. Alternately, with proper consideration of the periodic boundary conditions, the center of mass coordinate could be constrained to a steep potential well.

### 7.4.5 Summary

The technique of measuring diffusion at different finite box sizes and extrapolating to infinite volume is well known. But could a similar technique be used with different thermostat strengths? We saw above that the finite-size effect appears to be less pronounced with a stronger thermostat. Indeed, we should expect this, since the finite-size effect is due to momentum transfer around across the periodic boundary and acting back on a periodic image of the originating molecule. Since large $\gamma t_{\text{run}}$ decreases the effect of smaller boxes, we can presumably get away with running simulations in smaller boxes for shorter times by increasing $\gamma$, except that large values of $\gamma$ quickly cover up the true diffusion constant, as seen in (7.27). As a quick proof-of-concept, we try extrapolating (7.27) to $\gamma_{\text{t}} = 0$ using the smallest box size ($L = 2\text{nm}$) and the largest three thermostat values $\gamma_{\text{t}} \leq 0.01/\text{ps}$ (for which $\gamma_{\text{t}} t_{\text{run}} \gg 1$). Fitting to the line

$$\frac{1}{D(\gamma)} = \frac{1}{D_\infty} + a\gamma \tag{7.32}$$

we find $D_\infty = 0.53 \times 10^{-8} \text{m}^2/\text{s}$, about 10% larger than the measured NVE value at that box size, and 10% smaller than the accepted TIP3P diffusion constant [1].

With a greater understanding of the effects and interaction between box size and thermostat strength, we could presumably find the "sweet spot" to optimize the accuracy of the measured diffusion constant with the minimal computational resources. The next step in this direction is to redo the water simulations with longer run-times

($\gamma t_{\text{run}} \gtrsim 1$, cf. §7.4.2), randomized initial velocities (cf. §7.4.3), and greater care taken to cancel center of mass drift during the course of the simulation (cf. §7.4.4).

## BIBLIOGRAPHY


[1] P. Mark and L. Nilsson, "Molecular dynamics simulations of the ala-pro dipeptide in water: conformational dynamics of trans and cis isomers using different water models", *J. Phys. Chem. B* **105**, 8028 (2001).

[2] B. Dünweg, "Molecular dynamics algorithms and hydrodynamic screening", *J. Chem. Phys.* **99**, 6977 (1993).

[3] B. Dünweg and K. Kremer, "Molecular dynamics simulation of a polymer chain in solution", *J. Chem. Phys.* **99**, 6983 (1993).

[4] A. Malevanets and J. Yeomans, "Dynamics of short polymer chains in solution", *Europhys. Lett.* **52**, 231 (2000).

[5] I. C. Yeh and G. Hummer, "Diffusion and electrophoretic mobility of single-stranded RNA from molecular dynamics simulations", *Biophys. J.* **86**, 681 (2004).

[6] V. Lobaskin and B. Dünweg, "A new model for simulating colloidal dynamics", *New J. Phys.* **6**, 54 (2004).

[7] W. Jiang, J. Huang, Y. Wang, and M. Laradji, "Hydrodynamic interaction in polymer solutions simulated with dissipative particle dynamics", *J. Chem. Phys.* **126**, 044901 (2007).

[8] T. Soddemann, B. Dünweg, and K. Kremer, "Dissipative particle dynamics: A useful thermostat for equilibrium and nonequilibrium molecular dynamics simulations", *Phys. Rev. E* **68**, 046702 (2003).

[9] C. Pastorino, T. Kreer, M. Müller, and K. Binder, "Comparison of dissipative particle dynamics and Langevin thermostats for out-of-equilibrium simulations of polymeric systems", *Phys. Rev. E* **76**, 026706 (2007).