Equilibrium Shapes and Shape Transitions of Fluid Lipid-Bilayer Vesicles

by

Ling Miao

B.Sc., University of Science and Technology of China, 1984

M.Sc., Simon Fraser University, 1988

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
in the Department
of
Physics

SIMON FRASER UNIVERSITY
December 1992

APPROVAL

Name: Ling Miao

Degree: Doctor of Philosophy

Title of thesis: Equilibrium Shapes and Shape Transitions of Fluid Lipid-Bilayer Vesicles

Examining Committee: Dr. M. L. W. Thewalt
Chair

Dr. Michael Wortis
Senior Supervisor

Dr. David H. Boal

Dr. Michael Plischke

Dr. Barbara J. Frisken

Dr. Robijn Bruinsma
External Examiner
Professor, Department of Physics
The University of California, Los Angeles
California, U. S. A.

Date Approved: December 14, 1992

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

EQUILIBRIUM SHAPES AND SHAPE TRANSITIONS OF FLUID LIPID-BILAYER VESICLES

Author:

(signature)

LING MIAO

(name)

Dec. 17, 1992

(date)

ABSTRACT

The equilibrium shape of a fluid-lipid-bilayer vesicle is governed by the curvature elastic properties of the bilayer membrane. Experiments on artificially prepared giant vesicles have revealed a rich fauna of vesicle shapes and shape transformations, remarkably similar to what occurs in simple biological systems such as the human erythrocyte. The focus of this thesis work is to search for a theoretical understanding of these phenomena via the study of different models of lipid bilayers.

We first study the spontaneous-curvature (SC) model, which has two material parameters, the bending rigidity and the spontaneous curvature. The equilibrium shape, which minimizes the total bending energy at fixed surface area A and enclosed volume V, is explored over a significant range of the parameters, A and V. We map out for the first time a (shape) phase diagram, which shows the full systematics of equilibrium shapes and shape transformations predicted by the SC model. In particular, we show that both discontinuous budding (the eruption of a satellite connected to the parent body via a neck) and continuous vesiculation (the limiting case when the neck radius goes to zero) occur as equilibrium transitions in the appropriate regime.

Available experiments are not fully consistent with the predictions of either of the two extant membrane models (the SC model and the so-called area-difference or ΔA model). We propose that these discrepancies require generalization of the models to incorporate an area-difference elasticity (ADE), which allows the area difference between the two leaves of the bilayer to react elastically. We argue that this new term in the energy is comparable to those considered in the extant models. We introduce a new model (the ADE model) to incorporate this effect. We study budding and vesiculation in the ADE model. We find that, while vesiculation is still continuous, as it is in the SC model, the budding transition can occur either continuously or discontinuously, depending on the values of the elastic parameters and on the initial condition of the vesicle. For parameter values relevant to recent experiments, the ADE model predicts discontinuous budding. We discuss the connection between these predictions and laboratory observations.

ACKNOWLEDGEMENTS

My first and foremost thanks go to my thesis supervisor, Professor Michael Wortis, for suggesting this problem to me and introducing me to the field of soft-material physics. This thesis would not have been possible without his countless hours of discussion with me and his helpful advice. My general understanding of physics has certainly benefited from his clear and intuitive vision of many physics problems, which has been communicated pedagogically with patience. I feel very lucky to have been his student: As a supervisor, he always has time to discuss physics; as a friend, he is always caring and helpful. His conscientiousness, his uncompromising standards of doing research and his positive attitudes towards life in general have had important influence on me.

I am very grateful to Professor Michael Plischke, who has always been there for me when I needed his help and advice.

On the subject of this thesis, I have enjoyed many helpful and stimulating discussions with H-G. Döbereiner, B. Fourcade, J. H. Ipsen, M. Rao, U. Seifert, and R. K. P. Zia. I have also benefited from discussions with Professor E. Evans and Professor M. Bloom at the University of British Columbia. The technical help of M. Nikolić on the stability analysis is gratefully acknowledged.

During the last and most difficult phase of my thesis writing, I received much moral support and help from E. Levinson. My gratitude to him for his generosity and kindness is more than I can express with words.

I thank all my friends, who have made my times of happiness much more enjoyable and times of difficulty much more bearable. In particular, the friendship of M. Nikolić during the past year has always helped me to sustain the moments of hardship.

I am deeply in debt to my husband, Damin, for his unfailing support and understanding, and for all the sacrifices he has made in life over the past four years of this undertaking. His secretarial help during the last period of my writing has made it possible for me to have been able to graduate at the time of my preference.

Finally, financial support from Simon Fraser University and academic advice from my supervisory committee are gratefully acknowledged.

To my parents, who are on the other side of the Pacific Ocean.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Understanding red-blood-cell shapes has been a classic problem in cell biology since the invention of the microscope 300 hundred years ago. During the past two decades the distinctive discoidal contours of red blood cells at rest and the exquisite responses of these cells to diverse chemical perturbations in the form of a limited number of shape transformations (see Fig. 1.1) [1, 2] have attracted physicists' attention, as well [3]. An ultimate explanation, which relates these intriguing macroscopic properties of the red cell to the molecular structure of its cell membrane, is not to be reached easily. The reason is the following: The red-cell membrane has a marvelously complicated structure [4], as schematically pictured in Fig. 1.2. Its most elementary and indispensable structural component is a fluid bilayer formed from a mixture of cholesterol molecules and up to 1000 different kinds of lipids. Incorporated in this fluid-lipid-bilayer matrix are many diverse intramembrane proteins. Furthermore, the red-cell membrane includes a cytoskeleton, a cross-linked 2-dimensional protein network consisting principally of spectrin (a linear protein), which is anchored to the inner (cytoplasmic) surface of the lipid bilayer. This cytoskeletal network provides a strength and shear rigidity, which the fluid-lipid-bilayer component by itself would lack and which are crucial in maintaining red-cell integrity in the turbulent circulatory environment. It plays an important role in limiting red-cell shape deformations under conditions of high stress. Nevertheless, following the pioneering work of Canham, Helfrich and Deuling [6, 7, 8], the present consensus is that normal equilibrium red-cell shapes are principally (but not entirely) controlled by the energetics of the fluid-lipid-bilayer component.

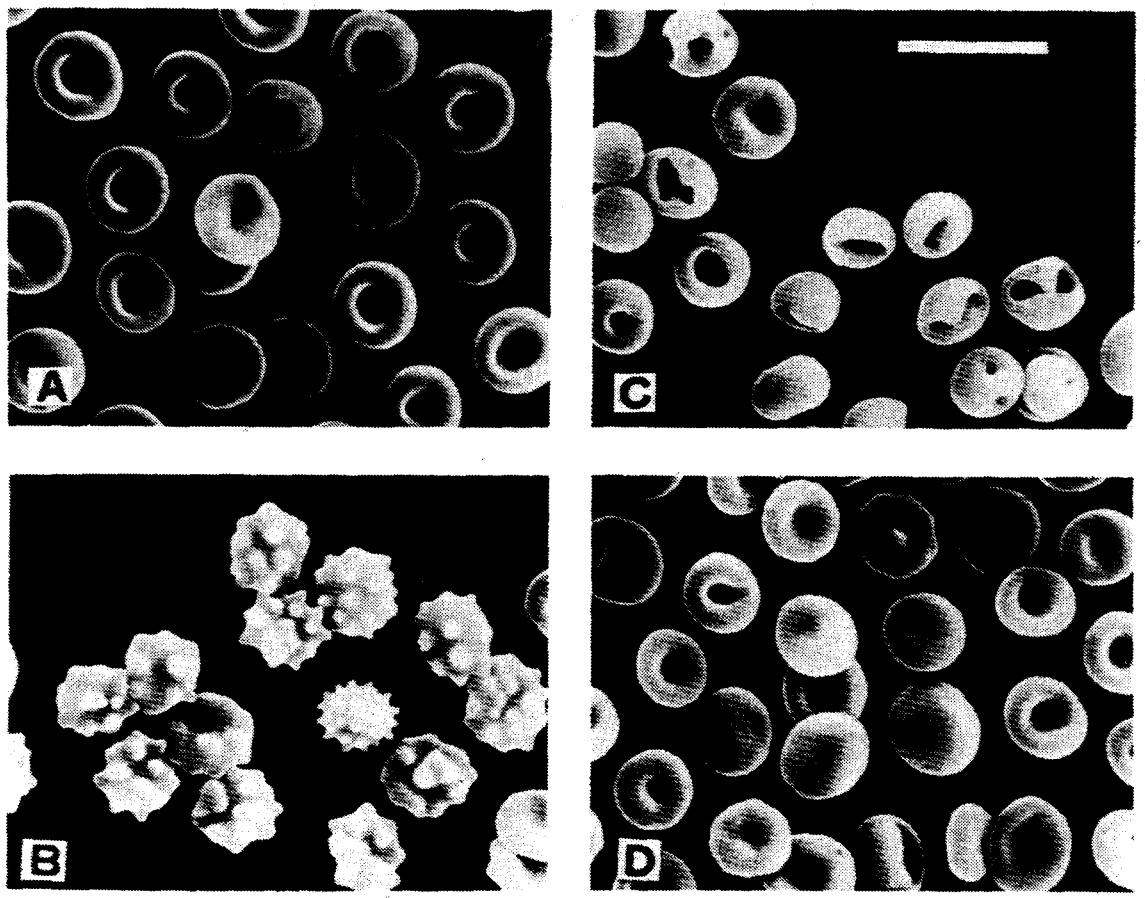


Figure 1.1: Shape transformations of erythrocytes, driven by external chemical perturbations (Ref. [2]). (A) A scanning electron micrograph of normal human erythrocytes in their discoidal shape (discocyte). (B) Cells treated with lysolecithin: They become spiculate echinocytes. (C) Cells with 37% of their cholesterol removed. This leads to invaginations and deepened dimples (stomatocytes). (D) Cells with reduced cholesterol, as in (C), were treated again with lysolecithin, as in (B). Note the restoration of the discoidal shape.

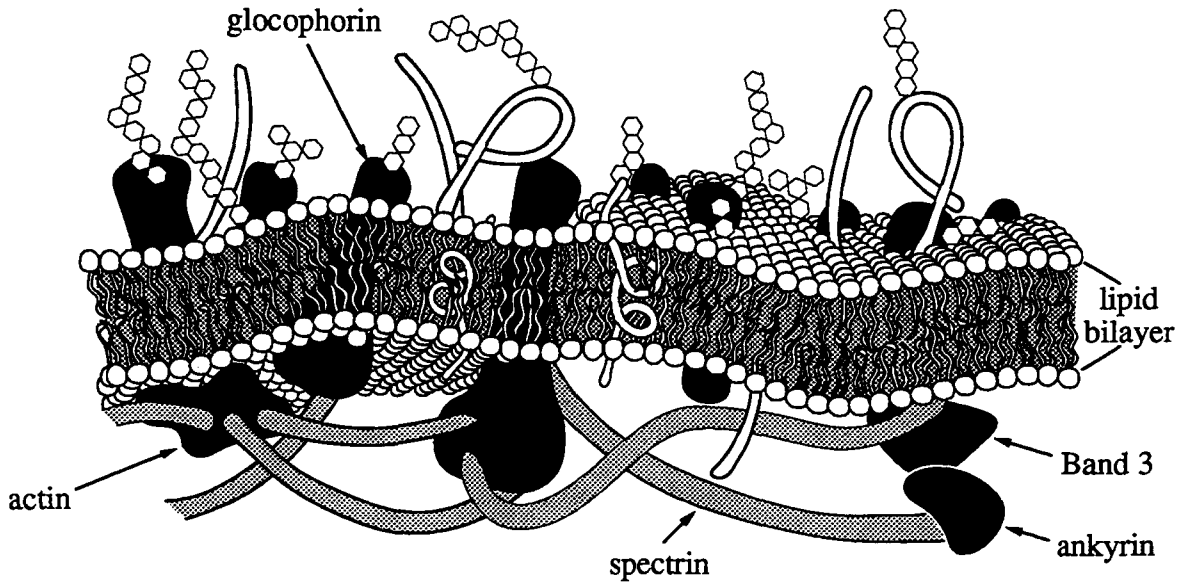


Figure 1.2: Anatomy of a red-cell membrane (Ref.[5]). The lipid bilayer is its main structural component. Band 3 and glycophorin are two types of intramembrane proteins. The cytoskeleton on the cytoplasmic (inner) side of the cell is formed primarily of linear spectrin molecules cross-linked via actin molecules. It is anchored to the lipid bilayer by its association with glycophorin via band 4.1 (protein), and with band 3 via ankyrin (protein).

It is of fundamental importance and interest to acquire a basic understanding of bilayer shapes and shape deformations and of the physical properties underlying such phenomena. Artificial lipid-bilayer vesicles serve as model systems ideal for this purpose. An artificial vesicle consists of a single encapsulating bilayer composed of one kind or at most a few kinds of lipids and free of complicating structures such as the cytoskeleton. In this thesis, we are concerned with the equilibrium shapes of lipid-bilayer vesicles, i.e., the geometrical surface profiles the vesicles assume when they are in equilibrium with their aqueous fluid environment. Although these model systems are enormously simpler in structure than red cells, they do exhibit a rich fauna of shapes and shape transformations, and there is a striking similarity between certain features of these physical model systems and the more complicated biological cells. It is tempting to think that these model systems are relevant in important ways to our understanding of elementary biological systems, such as red cells and of some elementary biological processes, such as endocytosis and exocytosis, which occur

frequently in real cells. On the other hand, we emphasize that the physics to be described in this thesis should be considered *physics done with biological materials* rather than true *biophysics*. The reasons are obvious: The membrane deformation in real biological cells is very complex and involves in general many types of molecules with highly non-trivial properties. For example, the budding-off of secretory vesicles from the plasma membrane of a cell is always associated with a particular type of protein molecules called Clathrins, which form a two-dimensional network-like coating over the budded vesicle. Furthermore, living biological cells are in general out of equilibrium. Hence, the biological implications of our understanding of model membranes are at best often indirect.

Our working hypotheses can be stated at the outset. We shall consider a lipid-bilayer vesicle freely suspended in an aqueous environment, e.g., distilled water or dilute sugar or salt solutions. There are two macroscopic length scales in this system: its linear size, which describes its shape and is typically 10–20 μm for laboratory vesicles studied [9], and the surface-persistence length, which describes its thermal undulations and is even larger than the linear size [10]. The thickness D of the bilayer membrane (typically 2–4 nm) is microscopic. Thus, the membrane is "thin" and can be represented on macroscopic length scales as a 2-dimensional geometrical surface. The equilibrium shape and, in general, the statistical behavior of such a fluctuating surface can be modelled by a hamiltonian which describes the elastic energy associated with curvature deformations of the membrane surface. The meaning of this statement will be made precise as we develop the fundamentals of the theory of vesicle equilibrium shapes in the remaining sections of this chapter and in the next chapter.

1.1 Physical properties of lipid-bilayer membranes

We first discuss some basic physical facts relevant to lipid bilayers and the orders of magnitude of important physical quantities [11] which are essential to the understanding of equilibrium vesicle shapes and shape transformations.

The elementary building blocks of artificial vesicles are glycerophospholipid molecules,

which are abundant in natural biological cell membranes. Giant (linear size of 10-20 μm) bilayer vesicles spontaneously form when glycerophospholipids are dispersed in water. The fact that phospholipid bilayers self-organize in this way is crucial to their occurrence as essential structural components of biological cell membranes and can be attributed to the dual (amphiphilic) nature of these molecules. They have structurally two well-defined parts: a phosphate-containing head group which is either polar or negatively charged and is, therefore, highly soluble in water (i.e., *hydrophilic*) and two fatty-acyl hydrocarbon chains that are attached to the headgroup and are almost completely insoluble in water (*hydrophobic*) (see Fig. 1.3(a)).

The hydrophobic feature is generic for nonpolar molecules or groups such as hydrocarbons, and it has highly non-trivial consequences [12]: When a non-polar particle is dissolved in water, the surrounding water molecules, otherwise isotropically arranged, must rearrange and become more ordered, an effect which decreases the entropy. The attraction between water molecules and the non-polar group is not strong enough to compensate for this entropy loss. The net result is an unfavourable free energy increase for a solution of non-polar particles in water. The hydrophobic effect induces a strong water-mediated attraction between non-polar particles or the non-polar parts of dissolved phospholipid molecules. It is this strong attraction that is the main driving force for the self-aggregation of phospholipids in water. In general, there can be several forms of phospholipid self-assembly. Examples are hexagonal structure, micelle structure, and bilayer structure, as shown in Fig. 1.3(b). For most two-chain phospholipids in the temperature and concentration ranges of interest, the favoured structure is a bilayer, in which the chains form a hydrocarbon interior and the heads face the water. We will only consider this structure in this thesis.

The lipid bilayer is self-bound and all the intermolecular forces must have reached a balance. The effect of such a balance is to reach an *optimal average area per lipid molecule*, a concept which will be extremely useful in developing the theory of the equilibrium vesicle shapes. Let us consider a lipid bilayer suspended in an aqueous solution, as sketched in Fig. 1.4. Headgroups are immersed in water. Separating the headgroup region and the hydrocarbon interior is an interface on each side of the bilayer, a substantial part of which

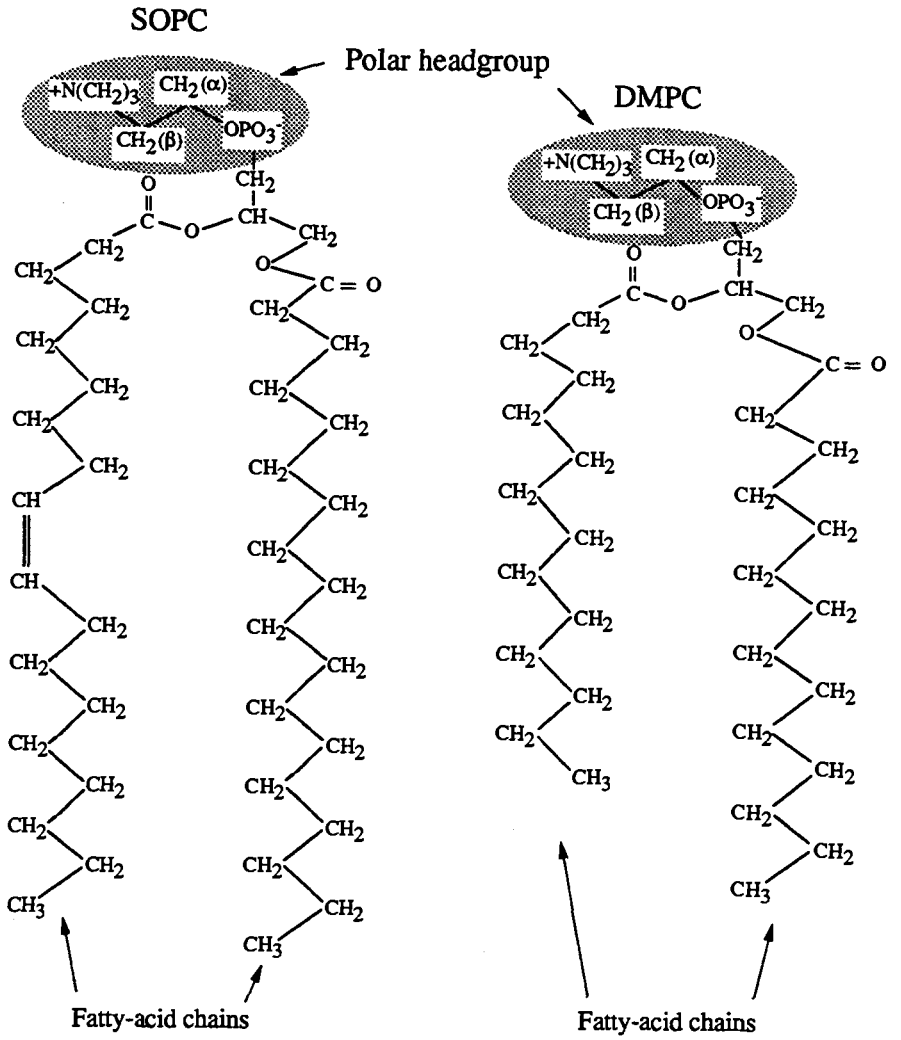


Figure 1.3: (a) Schematic illustration of molecular structures of SOPC and DMPC. The polar head group is phosphatidylcholine for both lipids. SOPC has one saturated and one unsaturated (note the double bond) fatty-acid chain; DMPC has two identical saturated fatty-acid chains.

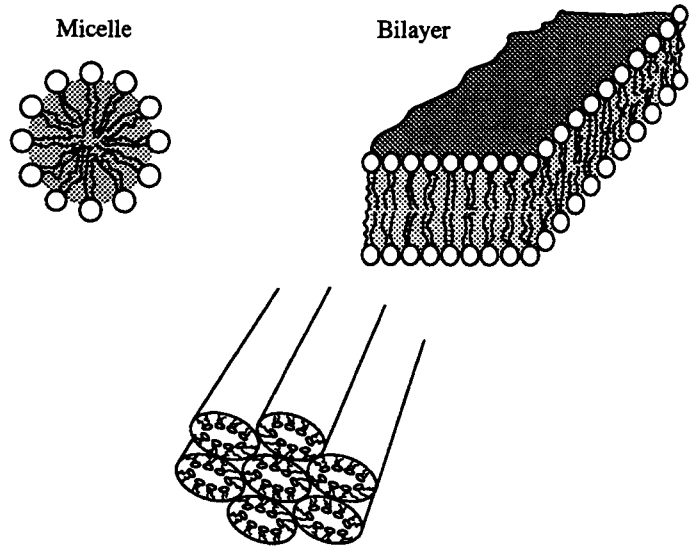


Figure 1.3: (b) Schematic diagram of examples of generic structures of lipid self-assembly in water. They are micelle structure, bilayer structure, and the inverted hexagonal structure.

is exposed to water-hydrocarbon contact[1]. Suppose that there are N molecules in a spatially uniform monolayer. Let a be the average local area per molecule and express the free energy of these N lipid molecules as a function of a,

$$F(a) = N\left[f_{pho}(a) + f_{phi}(a) + f_{int}(a)\right] , \tag{1.1}$$

where $f_{pho}(a)$ and $f_{phi}(a)$ correspond, respectively, to the effective attractive (hydro*pho*bic) and repulsive (hydro*phi*lic) interactions mediated by the water, while $f_{int}(a)$ represents the generically repulsive steric *int*eractions between the hydrocarbon chains. These various factors are not easy to treat exactly on a microscopic basis; however, an empirical approach proposed by Israelachvili [13] can at least provide an intuitive picture of the interplay between the opposing forces and give the correct order of magnitude for the relevant properties. The hydrophobic part can be modelled as,

$$f_{pho}(a) = \gamma(a - a_p) , \tag{1.2}$$

[1]Typically, the van der Waals type excluded area per headgroup is $0.4nm^2$, whereas the actual area per headgroup in a bilayer is $0.6nm^2$.

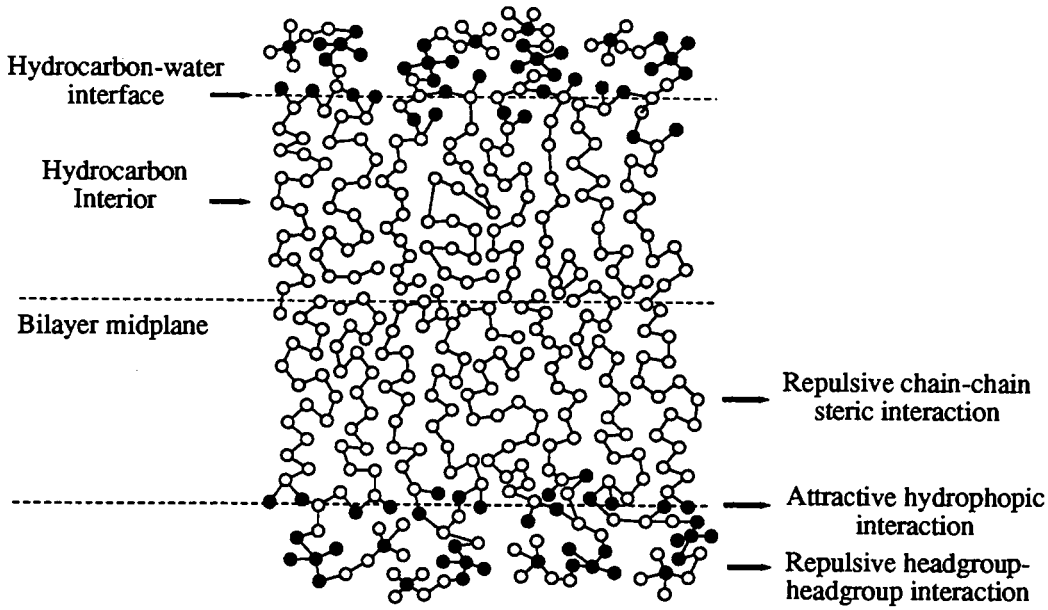


Figure 1.4: Schematic representation of the different kinds of interactions involved in a lipid bilayer (Ref.[13]). The effective repulsive forces arise from headgroup-headgroup and chain-chain interactions; the attractive forces originates from the hydrophobic effect. In general, these forces do not act in the same plane.

where $\gamma \approx 35 erg/cm^2$ is typical of the interfacial free-energy density of the water-hydrocarbon interface, and a_p is the excluded area per headgroup (of van der Waals type). The combined repulsive part varies roughly as a^{-1}, so

$$f_{phi}(a) + f_{int}(a) \cong C/a \,, \tag{1.3}$$

with $C \approx 1.15 \times 10^{-13} erg \cdot nm^2$.[2] The overall interfacial free energy per lipid molecule,

$$f(a) \equiv F(a)/N \cong \gamma(a - a_p) + C/a \,, \tag{1.4}$$

has a minimum at $a_0 = \sqrt{C/\gamma} \approx 0.57 nm^2$. One of the fundamental principles governing lipid self-assembly into bilayer structure is that, *in the absence of external forces, each lipid molecule in the bilayer will assume this optimal area at the water-hydrocarbon interface.*

[2]The surface pressure measurement of a lipid monolayer sitting at a water-hydrocarbon interface effectively measures the repulsive interaction. For di-acyl-phosphatidylcholine, such an empirical repulsive energy fits the measured data over a large range of surface area per molecule.

Expanding Eq. (1.4) for the monolayer free energy in the vicinity of this minimum leads to an approximate free energy per molecule for densities near a_0,

$$f(a) = f(a_0) + \frac{1}{2} 2\gamma (a - a_0)^2 / a_0 \, . \tag{1.5}$$

$2\gamma \cong 70 dyne/cm$ can be thought of as the area-compressibility modulus of the monolayer (roughly doubled for a bilayer). It is much smaller than a typical area-compressibility modulus of a solid thin shell of the same thickness. In other words, a lipid bilayer is much more deformable.[3] Indeed, phospholipid bilayer structures distinguish themselves from solid structures formed through strong intermolecular cohesion in this one important respect: The forces that hold phospholipid molecules together in the bilayer structure are not due to strong covalent or ionic bonds but arise from weaker van der Waals, hydrophobic, hydrogen-bonding and screened electrostatic interactions. Consequently, bilayers are soft and easily deformable.

From a physical point of view, lipid molecules in a bilayer have two different sets of fundamental degrees of freedom, internal conformations of their hydrocarbon chains and lateral translational motions. In general, a hydrated phospholipid bilayer can exhibit a variety of solid and fluid phases, depending, e.g., on temperature, which influences the interplay between the two sets of degrees of freedom. A typical low-temperature solid phase is the so-called *gel* phase, where the hydrocarbon chains are almost all in the *full-trans* (ordered) configuration and the molecules are packed into a two-dimensional hexagonal lattice. As temperature increases and passes a certain temperature called the *main transition temperature*, denoted T_m, the hydrocarbon chains become disordered, assuming many *gauche* configurations and transition to a fluid phase occurs. The main transition temperature T_m of a pure lipid bilayer depends upon parameters such as the length of the hydrocarbon chains and their degree of saturation (the number of C = C double bonds), and, therefore, it varies from one lipid to another. For example, T_m(DMPC) $\simeq$ 23°C and T_m(SOPC) $\simeq$ 5°C [14]. In this thesis, we restrict ourselves to the simplest fluid phase. Most of the experiments performed to investigate the shapes and shape transformations with which we shall

[3]The typical bulk modulus for a solid, e.g., aluminum, is $7.5 \times 10^{10} N/m^2$. An estimate for the area-compressibility modulus of a thin shell of thickness $2nm$ can be obtained by multiplying the bulk modulus by the thickness, a process which turns out to give $140 \times 10^3 dyne/cm$!

compare our theoretical study are done in temperature ranges well above the main transition temperature of the corresponding lipid bilayer. Bilayer fluidity greatly simplifies the analysis of the vesicle shape problem: *There is no shear present when a fluid-lipid-bilayer vesicle undergoes geometrical surface deformations!*

The lipid molecules in a *fluid* bilayer diffuse freely laterally in each individual monolayer, as molecules in an ordinary fluid do. The lateral diffusion coefficient is typically of order $10^{-8} - 10^{-7} cm^2/sec$ [15]. This means that a lipid molecule can move across the entire surface of a vesicle of μm size within a second!

In addition to their fast lateral diffusion, lipid molecules in a fluid bilayer are, in principle, capable of moving transversely from one monolayer to the other. However, such a motion–called flip-flop–is energetically very unfavourable. Indeed, in order for a lipid molecule to move intact from one leaf of the bilayer to the other, it would be necessary to drag the *hydrophilic* polar headgroup through the *hydrophobic* interior occupied by hydrocarbon chains. The transmembrane mobility of various phospholipids in both model membranes and biological membranes has been investigated. The results vary widely and are sometimes even controversial;[4] but, in general, it can be concluded that flip-flop rates are exceedingly slow in model membranes, with half-times ranging from hours to days [16]. In several of the biological membranes systems, however, a half-time of minutes or less has been reported [17]. The more rapid transmembrane diffusion may be due to specific proteins, such as "flipase", or to non-specific lipid-protein interactions which provide lower energy pathways for the migration of polar lipids across the membrane.

At the same time, a vesicle suspended in a fluid would in principle exchange lipid molecules with the adjacent fluid inside and outside. The low solubility of phospholipids in aqueous environments renders this lipid transport process insignificant. We arrive, therefore, at the following conclusion: The number of lipid molecules in each of the two individual monolayers of a closed vesicle can be considered to be separately conserved on the timescales of seconds–minutes necessary for mechanical equilibration.

[4]The primary reason for this seems to be that the conditions under which various experiments were performed are different and may affect lipid flip-flop in different ways. For example, an imbalance of pressure inside and outside a vesicle may induce significant lipid transport across the bilayer.

The last but not the least important physical property we would like to point out is the coupling of the two lipid monolayers that make up a bilayer. The monolayers are not bonded covalently, but are weakly adherent through van der Waals forces and, perhaps, slightly interdigitated [5]. Consequently, each individual monolayer is capable of collectively sliding relative to the other. As a result, the bilayer geometry does not impose any local constraint on the monolayer-surface-density fields.

1.2 Experiments on artificial vesicles

Although the main focus of this thesis will be the theoretical study of equilibrium shapes of lipid-bilayer vesicles, it is appropriate to review briefly the recent experiments performed mainly by two groups [9, 18, 19, 20] to investigate the equilibrium shapes and shape transitions of fluid-lipid-bilayer vesicles, for the following reasons: First, our theoretical work was originally motivated by the beautiful experimental observations of lipid-bilayer vesicle shapes recorded in the laboratory of Professor E. Evans at the University of British Columbia. Secondly, available experimental information is crucial to our theoretical study, since it provides both guidance and the ultimate touchstone of success.

Before describing the shapes which are observed, it will be helpful to explain how artificial vesicles are prepared in laboratory. Outlined below are the sequence of steps by which SOPC[5] vesicles are prepared in Professor Evans' lab. (i) A droplet containing the lipid dissolved in an organic solvent is placed on a roughened teflon disc. (ii) The solvent evaporates, leaving a residue of the pure dry lipid, presumed to exist in some lamellar phase. (iii) A stream of Ar saturated with water vapor is passed gently over the sample in order to swell the lipid layers as nondestructively as possible. The water apparently intercalates between the layers and lifts the lamellae apart. (iv) Liquid water is now introduced, producing a solution of vesicular structures. The mechanisms by which these structures rise off the substrate and close are not understood. (v) The solution at this stage includes, in addition

[5]SOPC is the abbreviation for the long name of one particular lipid, stearoyl-oleyl-phosphatidylcholine. It has a neutral head group. One of the two acyl chains (oleyl) is unsaturated while the other (stearoyl) is saturated.

to the "ideal" single-wall vesicles we shall discuss below, onion-skin structures with multiple walls, vesicles within vesicles, and multiple vesicles connected by small tubules. This "mess" is examined under the microscope. An "ideal" vesicle is selected and transferred to a second observation chamber by micropipette across an air barrier, thus effectively breaking any tubular connections. (vi) The selected vesicle (typically $10 - 20\mu m$ in diameter) is examined and recorded, usually by phase-contrast video microscopy. Vesicle-shape transitions are induced by adjusting the temperature of the experimental chamber and/or the osmolarity of the aqueous solution exterior to the vesicle.

The effects of increasing the temperature and changing osmolarity are the following: (a) Although, at given temperature T, the area A of the vesicle is fixed[6], increasing the temperature causes thermal expansion. Thus, $A = A(T)$ and increases as a function of temperature. (b) If there is an appreciable difference in osmolarity between the interior and the exterior of the vesicle, an osmotic pressure difference develops and water transport across the bilayer can and does occur. Thus, in the experiments, the interior and exterior osmolarities are carefully matched, so that negligible osmotic transport takes place on the timescales of seconds or minutes necessary for mechanical shape equilibration. The enclosed volume is kept fixed during shape equilibration. By controlling small osmotic differences on longer timescales, it is possible to inflate or deflate the vesicle at will. (c) The topology of the vesicles is normally spherical (however, toroidal vesicles have been seen [21]). (d) When the temperature is raised, the interior fluid is in principle subject to thermal expansion as well. It turns out, however, that this volume expansion is appreciably smaller than the area expansion discussed above,[7] so we may to a first approximation regard the volume as fixed in experiments where temperature is the only control parameter.

There follows now a brief catalog of some of the shapes and shape-transformation sequences seen in experiments by the two groups mentioned above:

A. Thermally induced discocyte-stomatocyte transitions [20, 23]. Discoid DMPC[8] vesicles

[6]The physical reason underlying this statement will be given in the next chapter.

[7]The volume expansivity of water is $3 \times 10^{-4}/K$ and the typical value of lipid bilayer area expansivity is $4 \times 10^{-3}/K$ [22].

[8]DMPC is the abbreviation for di-myristyl-phosphatidylcholine. It has the same head group as SOPC but its two acyl chains are identical and saturated.

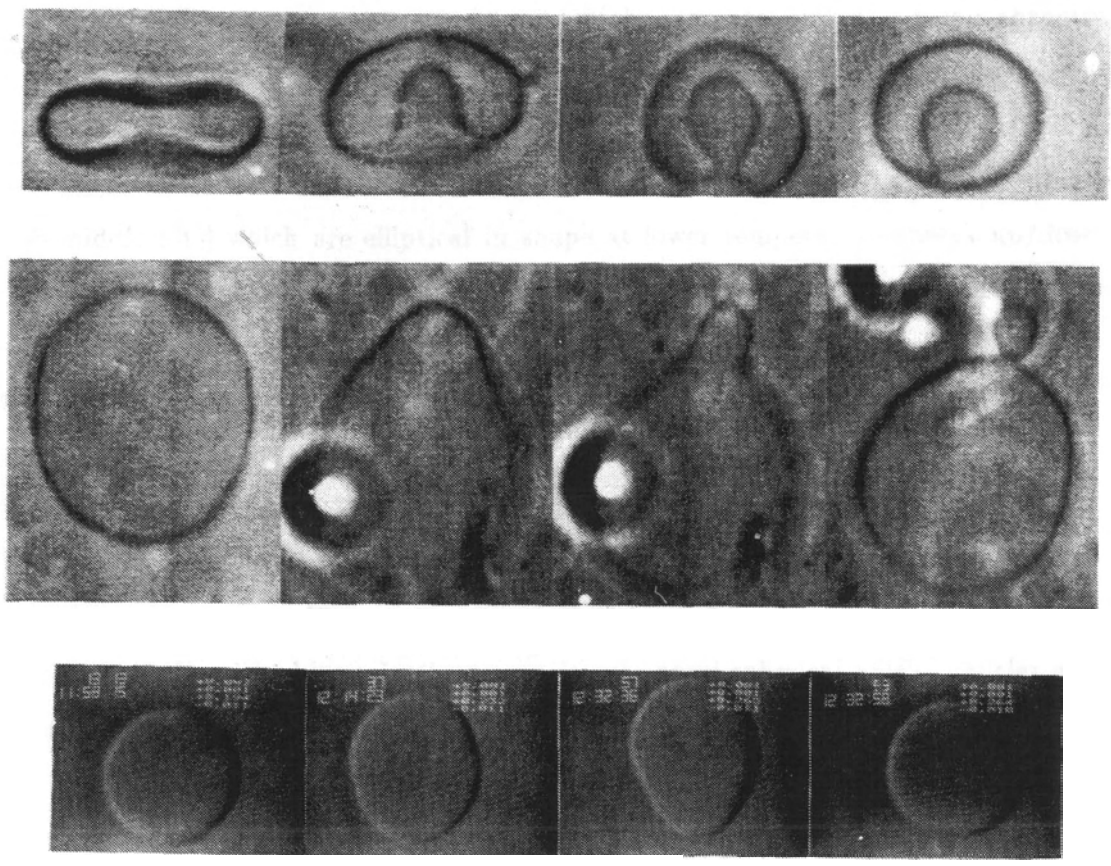


Figure 1.5: Overview of the experimental observations on artificially prepared giant vesicles. Top: Thermally induced discocyte-stomatocyte transition of DMPC vesicles: pictures were taken at 43.8, 43.9, 44.0 and 44.1°C (Ref. [23]). Middle: Thermally induced continuous budding of DMPC vesicles: the shapes were recorded at 31.4, 35.5, 35.6 and 35.8°C (Ref. [23]). Note that the pear-shaped intermediate is observed over the temperature range from 31.4 to 35.5°C, and is apparently the stable equilibrium shape in this range. Bottom: Thermally induced discontinuous budding of SOPC vesicles (by courtesy of H-G Döbereiner). The pear shape is observed only as a dynamical transient. Note that the time for the pear shape to become vesiculated or nearly vesiculated is very short (20 seconds) (see the time shown on the upper-left conner of each picture). This transition is first order.

change to stomatocytes as the temperature is raised. Up/down symmetry is broken at this transition. The shape apparently evolves continuously. The transformation is analogous to that shown in Fig. 1.1 for red cells. At some higher temperature, the shape abruptly becomes invaginated. This process is illustrated in Fig. 1.5(top set), following the original which appeared in Ref. [23].

B. Thermally induced *continuous* budding via "pear" shapes [20, 23]. DMPC vesicles [Fig. 1.5(middle set)] which are elliptical in shape at lower temperatures break up/down symmetry to become pear-shaped as T is increased. The pear shape is observed over a few degrees of temperature increase and *suddenly* jumps at a higher temperature to a seemingly vesiculated shape which consists of two spheres connected through an invisible (possibly microscopic) neck.

C. Thermally induced *discontinuous* budding [9]. SOPC vesicles go through a temperature-driven and apparently discontinuous budding as T is increased from a prolate ellipsoid to an almost vesiculated shape [Fig. 1.5(bottom set)] without the pear-shaped intermediates.[9]

D. Osmotically induced budding. In these experiments quasi-spherical SOPC vesicles were produced in 200 mOsm sucrose solution, then pipetted into 200 mOsm glucose, and finally transferred to 240 mOsm glucose at all fixed T. This leads to osmotic deflation of the vesicle, which was allowed to proceed at fixed temperature. Budding transition to a narrow-necked or vesiculated shape was observed.

1.3 Overview of the thesis

It is well recognized that a fluid-lipid-bilayer surface has elastic resistance against curvature (or bending) deformations. Several versions of elastic curvature energies have been

[9] Actually, the present status concerning the experiments is not entirely clear: The distinction between the characteristics of the budding transitions described in Sequences B and C has not been made as unambiguously as stated. The identification of the symmetry-breaking shape evolution as a continuous equilibrium transition, as reported in Case B, is based on the fact that the asymmetric pears are observed over several degrees of temperature increase, which was implemented on the time scale of minutes. As we shall discuss in Sec. 4.4.1, the subtle interplay between this time scale and the typical relaxation time of overdamped shape fluctuations may allow one to have a different interpretation of the observation. By the same token, the claim for discontinuous budding, as stated in Sequence C, should be taken with caution too, although we are inclined to think that it is the case, based on more recent experimental observations [24].

proposed to model such resistance. In Chapter 2, we introduce three curvature models, the area-difference-elasticity (ADE) model, the spontaneous-curvature (SC) model and the area-difference (ΔA) model. The emphasis is on the ADE model: We present in detail a "microscopic" derivation of its curvature energy, by incorporating the fundamental physical properties of a fluid-lipid-bilayer vesicle, as reviewed in Chapter 1. The ADE model not only takes into account the elastic effect associated with changes in local curvatures on the vesicle surface, but also allows the area difference of the two leaves of the bilayer to react elastically (the ADE effect). It is, therefore, a generalization of the two extant models, the SC model and the ΔA model which are from this perspective limiting cases of the ADE model.

Chapter 3 describes our systematic investigation of the taxonomy of equilibrium shapes and shape transformations within the context of the SC model. We first develop a general framework, applicable to all the vesicle-shape models. To this end, we introduce the constrained minimization of the model free-energy functional, the Euler-Lagrange differential equations to which it leads, the regular Euler shapes as solutions to the Euler equations, and the resulting (shape) phase diagram. Several classes of stationary shapes and shape transitions are found and examined in detail. In particular, the importance of multiplets, configurations in which two or more regular Euler shapes join each other tangentially and possess collectively the total surface area and volume, in competition for the lowest-energy shape is recognized for the first time. Close associates of multiplets, shapes consisting of several segments connected via small open necks, are shown to undergo "vesiculation", i.e., to become multiplets as the necks are driven to zero by the changes in control parameter. We derive a general criterion for the location of vesiculation boundaries. We also identify "budding" as an equilibrium shape transition, at which a parent vesicle (usually quasi-spherical) erupts to form one or more small daughter vesicles, bound to it through small necks. In addition, we obtain a set of useful results for spheres, sphere multiplets and nearly spherical shapes. The analytical treatments of the mathematical bifurcation structures both near vesiculation and close to spheres are relegated to Appendix C and Appendix D, respectively. Finally, we map out the full phase diagram through numerical

and quasi-analytical means. This phase diagram summarizes the specifics of equilibrium shapes and shape transitions in the SC model.

Neither the SC model nor the ΔA model gives a fully consistent explanation of all the results observed in controlled shape experiments. Chapter 4 reports our study of the more general ADE model. This research was motivated by the conjecture that the incorporation of the area-difference-elasticity effects in the ADE model may lead to a better understanding of the experiments. We focus on the two prominent shape transitions, the budding transition and the vesiculation transition. We find that the budding transition can occur either continuously or discontinuously, depending on the values of the elastic parameters and the initial conditions of the vesicle. For parameter values relevant to recent experiments, the ADE model predicts discontinuous budding. We discuss the connections between these predictions and laboratory observations.

Chapter 2

Continuum Models for Fluid Lipid-bilayer Vesicles

Over the past two decades, several models have been developed to describe the shape of a single isolated vesicle formed by a fluid lipid bilayer. They are all continuum models, valid in the limit where typical radii of curvature of the vesicle are large on the scale of the bilayer thickness. Among them are two now-traditional ones, the spontaneous-curvature model (which we shall denote hereinafter as the SC model) and the area-difference or ΔA model, definitions of which will be given later in this chapter. These models were originally studied [8, 25, 26] in order to explain the observed shapes of red blood cells. Indeed, twenty years ago Helfrich and Deuling [8] showed that the SC model contains a rich catalogue of possible vesicle shapes with many similarities to those observed for real red blood cells (Fig. 1.1). It soon became clear, however, that the spontaneous curvature parameter in the SC model would require further explanation. In particular, all vesicles in the same chemical preparation do not share the same spontaneous curvature. The ΔA model [23, 25, 26, 27] provided an escape from this difficulty by linking (as we shall see below) the closure process for each individual vesicle to the effective values of its spontaneous curvature. The full systematics of equilibrium shapes and shape transformations predicted by these two models has only been revealed recently through our work[1] [28] and the parallel work of Seifert, Berndl and Lipowsky [27].

[1] Chapter 3 contains an account of our work on the systematics of the SC model.

Only recently has it been possible to compare the predictions of the SC and the ΔA models with the observed results of controlled experiments on the artificial lipid-bilayer vesicles, as described in Chapter 1. It has turned out [29] (also see Chapter 4) that considerable discrepancy persists between the predictions of the two models and the experimental findings. In particular, neither the SC model nor the ΔA model can be made consistent with all observations. This disagreements has motivated reflection on the physical basis of the two models. Stimulated by the earlier remarks of Helfrich [7], Evans [30] and Svetina and Žekš [31], we have concluded that there exists in the problem an additional physical parameter, which is not adequately represented in either the SC model or the ΔA model. This new parameter, an area-difference elasticity, provides an elastic correction to the area difference of the ΔA model. On the basis of crude estimates (see later text in this chapter), we believe that these elastic effects are fully comparable in magnitude to the effects contained in the SC and the ΔA models. The study of a new model (The ADE model) incorporating this area-difference-elasticity effect forms Chapter 4 of this thesis.

This chapter is devoted to the introduction of the three continuum models, the SC model, the ΔA model and the ADE model. This discussion continues the development of fundamentals started in Chapter 1. It will also provide a comprehensive review of both our present best understanding of the physics of fluid-lipid-bilayer vesicles and the historical development in this field. For pedagogical reasons we will not follow the course of history. Rather, we will start by presenting the more general ADE model. Both the SC model and the ΔA model will then be introduced as special cases, in which, as will become clear, the ADE effect is either overlooked or overestimated.

2.1 A general curvature-energy model–the ADE model

We consider a single giant artificial lipid-bilayer vesicle suspended in an aqueous solvent. In the continuum limit, the global geometrical quantities characterizing its shape S are its surface area A and enclosed volume V. The basic local geometrical variables are the two local principal curvatures C_1 and C_2 at each point on the vesicle surface. The pioneering

work of Canham, Helfrich, Deuling and Helfrich, Evans and Skalak [6, 7, 8, 32] has led to the belief that the shape of such a vesicle is governed primarily by the elastic (or mechanical) properties of the encapsulating lipid-bilayer membrane [6, 7, 32]. Each continuum model proposed to describe this macroscopic (typically μm scale) phenomenon, therefore, consists of an elastic-energy functional depending on the shape S of the vesicle, and a set of constraints on the global geometrical quantities. The global constraints reflect the physical conditions which the vesicle is assumed to be under.

The ADE model has the following energy functional,

$$W[S] \equiv \frac{\kappa}{2} \oint dA[C_1(\mathbf{r}) + C_2(\mathbf{r}) - C_0]^2 + \frac{\bar{\kappa}}{2}\frac{\pi}{AD^2}(\Delta A - \Delta A_0)^2 \,, \tag{2.1}$$

where the integral runs over the surface of the vesicle. The local principal curvatures $C_1(\mathbf{r})$ and $C_2(\mathbf{r})$ are defined to be positive when the center of curvature is towards the vesicle interior. C_0 is a phenomenological parameter called the *spontaneous curvature*, which we shall discuss below. κ is the *local* bending elastic constant of the bilayer membrane and has the dimension of energy. The second term in the energy functional, characterized by another bending elastic constant $\bar{\kappa}$, is called the *non-local* bending energy and derives from the fact that the vesicle surface is composed of two monolayers, with (possibly) different compositions, which can and may respond differently when the vesicle surface undergoes a geometrical deformation. Both elastic constants κ and $\bar{\kappa}$ are positive. D, as defined before, is the membrane thickness. ΔA_0 represents the *preferential* surface-area difference between the two monolayers. In the simplest case, where the bilayer contains only a single species of lipids with an optimal area a_0 per molecule, $\Delta A_0 \equiv (N^{out} - N^{in})a_0$, with N^{out} and N^{in} being the number of molecules in the outer and inner monolayers, respectively. ΔA is the actual *geometrical* area difference, another *global* quantity dependent on the specific vesicle shape: If the thickness D is regarded as a constant over the whole membrane, then it is an elementary geometrical result that ΔA is related to the integral of the local mean curvature $C_1(\mathbf{r}) + C_2(\mathbf{r})$ over the vesicle surface,

$$\Delta A = D \oint dA(C_1 + C_2)[1 + O(D/R)] \,, \tag{2.2}$$

where R is the characteristic radius of curvature of the vesicle. The model (2.1) will be

referred to as the ADE model (ADE stands for the "area difference elasticity"), because it involves the (unconventional) elastic energy connected with the monolayer area difference.

The remainder of this section will be devoted to a "microscopic" derivation of the ADE energy functional (2.1). A precise understanding of the macroscopic elastic properties represented by the energy functional (2.1) at the molecular level can be expected to be prohibitively difficult, because of the complexity of the molecular interactions involved. There are interesting attempts at developing fully microscopic theories along this line [33, 34]; however, we shall content ourselves in this thesis with a simple, semi-phenomenological derivation, which gives important insights into the essential and relevant physics embedded in the energy functional (2.1).

We begin by looking at one individual monolayer component of the bilayer membrane, assumed to be in its fluid state. We recall from Chapter 1 the concept of the optimal area per molecule a_0 and the quadratic expansion of the elastic energy per molecule of a fluid lipid-monolayer, Eq. (1.5),

$$f(a) = f(a_0) + \frac{K}{2} a_0 (a/a_0 - 1)^2 \,. \tag{2.3}$$

In the considerations there, all the interactions (i.e., the attractive hydrophobic interaction, the repulsive hydration force between the headgroups, the repulsive steric chain-chain interaction, and less importantly the van der Waals interactions), are assumed to act in the same plane. A more realistic picture of these interactions is depicted in Fig. 2.1, in which we assume that the centers of action of the opposing forces are localized in two regions separated by a distance δ, one corresponding to the heads of the molecules, the other, to the chains. The elastic energy per molecule can then be written as

$$f(a) = \frac{1}{2} K_h a_{h,0} (a_h/a_{h,0} - 1)^2 + \frac{1}{2} K_c a_{c,0} (a_c/a_{c,0} - 1)^2 \,, \tag{2.4}$$

where a_h and a_c are the *actual* area per molecule at the level of heads and chains, respectively. $a_{h,0}$ and $a_{c,0}$ are the *optimal* area per molecule and may in general be different. K_h and K_c are elastic area-stretching constants (or area-compressibility moduli). It may be helpful to picture the shape of a single lipid molecule as a truncated cone. The energy (2.4)

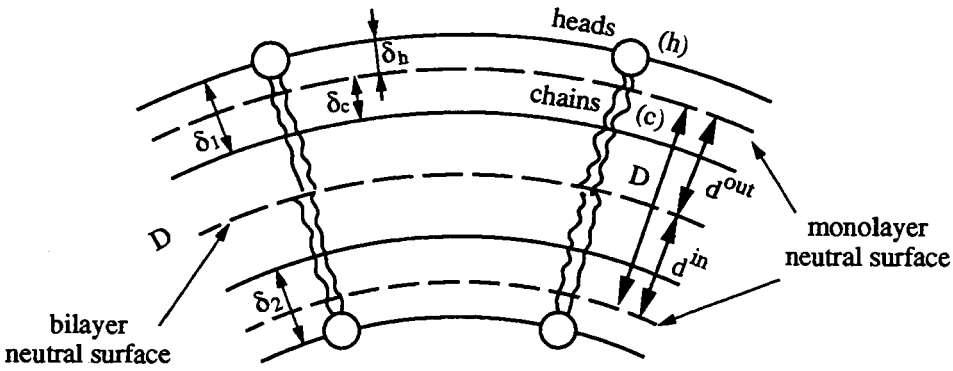


Figure 2.1: A simple microscopic model of the curvature elasticities of a lipid-bilayer. It is assumed that different forces are localized in a discrete set of surfaces, represented by solid lines. Various neutral surfaces are shown with the dashed lines. In the monolayer discussion, we consider a single layer of lipid molecules. The distance $\delta = \delta_h + \delta_c$ separates the surfaces of head-head and chain-chain interaction.

has a simple mechanical interpretation: $(a_h/a_{h,0} - 1)$ and $(a_c/a_{c,0} - 1)$ can be thought of as the local, isotropic elastic strains representing area dilation/compression. The corresponding elastic surface tensions (or stresses) are, then, $K_h\,(a_h/a_{h,0} - 1)$ and $K_c\,(a_c/a_{c,0} - 1)$, respectively. Note that the stresses are isotropic, as is characteristic of a fluid surface.

It is necessary now to choose a reference surface, with respect to which the geometrical characteristics of the monolayer surface at the macroscopic scale (such as surface area and local curvatures) can be defined. We will make a particular choice of such a reference surface below and will call it the "neutral surface". Let us call a the area per molecule at this neutral surface, and δ_h, δ_c, the respective distances of the head and the chain surfaces from the monolayer neutral surface. It follows from simple geometry that

$$\begin{aligned} a_h &= a\,[1 + \delta_h(C_1 + C_2) + \delta_h^2 C_1 C_2 + O((\delta_h C)^3)] \\ a_c &= a\,[1 - \delta_c(C_1 + C_2) + \delta_c^2 C_1 C_2 + O((\delta_c C)^3)]\,, \end{aligned} \tag{2.5}$$

with $\delta = \delta_h + \delta_c$. Two important implications of Eq. (2.5) should be pointed out: (i) If the monolayer is bent, as represented by non-zero local curvatures, the area change at the level of the headgroups is different from that at the level of the chains; (ii) On the other hand, the changes are constrained locally, since the head and tail surfaces are rigidly connected

through molecules. The local constraint is

$$a_h = a_c\,[1 + \delta(C_1 + C_2) + O(\delta^2 C^2)]\,. \tag{2.6}$$

By introducing Eq. (2.5) into the elastic energy (2.4), one obtains

$$\begin{aligned} f(a) \;=\; & \frac{1}{2}\left\{ \left(\frac{K_h}{a_{h,0}} + \frac{K_c}{a_{c,0}} \right) a^2 - 2(K_h + K_c)a \right. \\ & + \left(\frac{K_h \delta_h^2}{a_{h,0}} + \frac{K_c \delta_c^2}{a_{c,0}} \right) a^2 (C_1 + C_2)^2 \\ & + 2\left[\left(\frac{K_h \delta_h}{a_{h,0}} - \frac{K_c \delta_c}{a_{c,0}} \right) a^2 - (K_h \delta_h - K_c \delta_c)\, a \right] (C_1 + C_2) \\ & \left. + 2\left[\left(\frac{K_h \delta_h^2}{a_{h,0}} + \frac{K_c \delta_c^2}{a_{c,0}} \right) a^2 - \left(K_h \delta_h^2 + K_c \delta_c^2 \right) a \right] C_1 C_2 \right\}\,. \end{aligned} \tag{2.7}$$

By choosing the position of the neutral surface so that[2]

$$\frac{K_h \delta_h}{a_{h,0}} - \frac{K_c \delta_c}{a_{c,0}} = 0\,, \tag{2.8}$$

we arrive at the following expression for the energy density (energy per unit area)[3]:

$$\begin{aligned} f/a_0 = f_0/a_0 \;+\; & \left\{ \frac{K}{2}(a/a_0 - 1)^2 + \frac{\kappa}{2}[(C_1 + C_2) - C_0]^2 \right. \\ & \left. - \frac{\kappa}{2} 2C_0 (a/a_0 - 1)(C_1 + C_2) + \kappa_g C_1 C_2 \right\}\,, \end{aligned} \tag{2.9}$$

where the relevant parameters are given by

$$a_0 \equiv \frac{(K_h + K_c) a_{h,0} a_{c,0}}{K_h a_{c,0} + K_c a_{h,0}}\,, \tag{2.10}$$

$$K \equiv K_h + K_c\,, \tag{2.11}$$

$$\kappa \equiv \frac{K_h K_c (K_h + K_c) a_{h,0} a_{c,0}}{(K_h a_{c,0} + K_c a_{h,0})^2}\, \delta^2\,, \tag{2.12}$$

$$C_0 \equiv \frac{1}{\delta} \frac{(a_{h,0} - a_{c,0})(K_h a_{c,0} + K_c a_{h,0})}{(K_h + K_c) a_{h,0} a_{c,0}}\,, \tag{2.13}$$

[2]Note that this eliminates the coupling term in Eq. (2.7), $a^2(C_1 + C_2)$.

[3]This derivation is, in fact, just an expansion of the elastic energy in terms of the small quantities $(a/a_0 - 1)$ and (δC). Only quadratic terms in these small quantities are kept in the final expression.

and

$$\kappa_g \equiv \frac{K_h K_c (K_h a_{c,0} - K_c a_{h,0})(a_{h,0} - a_{c,0})}{(K_h a_{c,0} + K_c a_{h,0})^2} \delta^2 . \tag{2.14}$$

The choice of the neutral surface given by Eq. (2.8) has the following interpretation: The local elastic energy density corresponding to an arbitrary deformation of the monolayer can be decomposed into one contribution arising from the area dilation of the neutral surface, represented by $(a/a_0 - 1)$, and another due to the bending deformation, described in terms of the two principal curvatures. Under this convention a *pure bending* is defined to be a deformation which does not cause any area variation of the neutral surface and leads only to change in the curvatures. In this case the only contribution to the elastic energy density comes from the curvature energy.

This derivation can already provide several insights into the physics contained in the part of the curvature-related elastic energy in Eq. (2.9). Eq. (2.12) tells us that the local bending elastic constant of the monolayer $\kappa \propto K_{h,c}\delta^2$, (not δ^3 as it would be for a solid thin shell [35]) and that this elastic constant is always positive. A crude estimate of this quantity can be obtained easily: For typical values, $K_h = K_c = K/2 \approx 35 erg/cm^2$ (see Chapter 1) and $\delta \approx 2nm$, we find $\kappa \sim 5 \times 10^{-13} erg$, a number which is in good agreement with the experimentally measured value [22]. The physical origin of the monolayer spontaneous curvature C_0 defined in Eq. (2.13) becomes intuitively clear. It is related to the steric asymmetry of the molecular shape. Since the steric shape is a phenomenological representation of all the microscopic interactions, one can imagine controlling this asymmetry by modifying some of the interactions, e.g., the interaction between headgroups, which intimately depends on the chemical or electrostatic conditions in the aqueous environment adjacent to the monolayer. In contrast to the bending rigidity κ, the elastic constant κ_g associated with the local Gaussian curvature $(C_1 C_2)$ can be positive or negative, depending upon the relative magnitudes of the head and chain contributions to Eq. (2.13). This fact is important for soluble amphiphiles, since varying the sign of κ_g of a monolayer can influence the phase behavior of the amphiphile-water system [36]. Eq. (2.14) also suggests that κ_g may be smaller than κ by a couple of orders of magnitude due to its dependence on the factor $[(K_h a_{c,0} - K_c a_{h,0})/(K_h a_{c,0} + K_c a_{h,0})](a_{h,0}/a_{c,0} - 1)$, which can be small for lipid

molecules with only slight asymmetry in shape.

The local elastic energy density f/a_0 given in Eq. (2.9) is the simplest Landau energy involving the two sets of coarse-grained degrees of freedom fundamental to a lipid monolayer, the "internal" degrees of freedom, represented by $(a/a_0 - 1)$ and corresponding to the translational motion of lipid molecules within the monolayer, and the "external" (or shape-dependent) degrees of freedom, such as local curvatures, connected with the shape of the monolayer surface. This Landau functional contains all possible local, isotropic, and Euclidean-invariant (i.e., invariant under translations and rotations in the 3-dimensional embedding space) terms through second order in $(a/a_0 - 1)$ and in the local curvatures, including the coupling term between the two sets of degrees of freedom: Up through second derivatives of the shape, the only local Euclidean invariants are the principal curvatures C_1, C_2. If we require isotropy, then C_1 and C_2 must enter symmetrically. Thus, at first order in C, there is only one invariant, the local mean curvature $H = C_1 + C_2$, while at second order there are two, H^2 and the Gaussian curvature (C_1C_2). Additional terms, involving higher orders of $(a/a_0 - 1)$ and C_1, C_2 and higher derivatives of the shape, are in principle present but neglected herein (see Helfrich's recent work for the effect of these higher-order terms [37]). The coupling term between the "internal" and "external" degrees of freedom in Eq. (2.9) will also be neglected, since it effectively makes a small contribution of order $(a/a_0 - 1)$ to the spontaneous curvature C_0.

Another important piece of physics revealed by this derivation should be emphasized: There are two (we neglect the Gaussian curvature term for the moment) vastly different energy scales associated with the monolayer deformation, a fact which can be easily seen if we write the total elastic free energy of the monolayer approximately in the form,

$$F(N) \equiv \int dN f \sim \frac{KA_0}{2}(A/A_0 - 1)^2 + \frac{\kappa}{2}[A_0(C_1 + C_2 - C_0)^2] \,, \tag{2.15}$$

where A_0 is the surface area of the whole monolayer. The energy scale (KA_0) is connected with surface-area dilation. The ratio of this energy to the bending energy constant κ is roughly $A_0/\delta^2 \sim (R/\delta)^2 \sim 10^6$. An area dilation which costs an energy comparable to the bending energy is, therefore, only of order $\delta/R \sim 10^{-3}$.

Having discussed the physics of a single monolayer, we turn now to the properties of a closed membrane bilayer composed of two such monolayers. It will be assumed that the thickness D of the bilayer, or, more specifically, the separation between the monolayer neutral surfaces, stays constant. The elastic energy of the lipid bilayer can thus be written as,

$$W \equiv \oint_{A^{out}} \frac{dA^{out}}{a^{out}} f^{out} + \oint_{A^{in}} \frac{dA^{in}}{a^{in}} f^{in} , \tag{2.16}$$

where f^i has the form given in Eq. (2.9),

$$f^i \equiv a_0^i \left\{ \frac{1}{2} K^i (a^i/a_0^i - 1)^2 + \frac{1}{2}\kappa^i [(C_1^i + C_2^i) - C_0^i]^2 + \kappa_g^i C_1^i C_2^i \right\} , \tag{2.17}$$

and the index i ($i = out, in$) labels the individual monolayers.[4]

In this approximation, the surface density variables, a^{out} and a^{in}, can always be equilibrated, independent of local curvature changes. Indeed, as noted in Chapter 1, there is no *local* constraint imposed on a^{out} and a^{in}. Consequently, both a^{out} and a^{in} will adjust themselves independently of one another to minimize the total elastic free energy. The only constraints are that the number of molecules in each monolayer is conserved separately. The constrained minimization of the elastic energy with respect to a^i can then be formulated as follows:

$$\begin{aligned} \frac{\delta}{\delta a^{out}} \left(W - \mu^{out} \oint_{A^{out}} \frac{dA^{out}}{a^{out}} - \mu^{in} \oint_{A^{in}} \frac{dA^{in}}{a^{in}} \right) &= 0 \\ \frac{\delta}{\delta a^{in}} \left(W - \mu^{out} \oint_{A^{out}} \frac{dA^{out}}{a^{out}} - \mu^{in} \oint_{A^{in}} \frac{dA^{in}}{a^{in}} \right) &= 0 , \end{aligned} \tag{2.18}$$

where the chemical potentials μ^{out} and μ^{in} are used to enforce the constraints on the number of molecules, $\oint_{A^i} dA^i/a^i = N^i$ in each monolayer i. The two variations (2.18) lead to two local equations,

$$\left(\frac{a^i}{a_0^i} - 1 \right) + \frac{\mu^i}{K^i a^i} - \frac{f^i}{K^i a^i} = 0 . \tag{2.19}$$

[4]Two remarks are in order here. First, in writing out the total energy of the bilayer, as is in Eq. (2.16), we are performing a mean-field treatment of the system, in which certain degrees of freedom have been integrated out. Secondly, we have neglected the van der Waals interactions between molecules within the two leaves. One might expect that such interactions would induce correlated inhomogeneity in the surface density distribution of each monolayer. However, it is easy to show that at the mean-field level solutions with uniform lateral density continue to exist and, for these solutions, the effect of the van der Waals interaction is merely to add an overall constant to (2.24).

In Eq. (2.19) the only term depending on the local curvatures (included in f^i) comes from the curvature-energy contribution, as given in Eq. (2.17). It describes a "secondary" effect (of the order of $(\delta/R)^2$), and can be neglected. It follows from Eq. (2.19) that the density deviation at equilibrium, $(a^i - a_0^i)$, is independent of position on the monolayer surface. Thus, the area dilation of each monolayer is a *global* quantity, characteristic of each monolayer surface, and is equal to $(A^i/A_0^i - 1)$. We thus rewrite the bilayer energy (2.16),

$$W = \sum_i \left\{ \frac{K^i A_0^i}{2}(A^i/A_0^i - 1)^2 + \oint_{A^i} dA^i \left[\frac{\kappa^i}{2}[(C_1^i + C_2^i) - C_0^i]^2 + \kappa_g^i C_1^i C_2^i \right] \right\}$$
$$+ \text{ higher-order corrections} . \tag{2.20}$$

A neutral surface for the bilayer can be defined in the same way that was done for the monolayer,

$$\frac{K^{out} d^{out}}{A_0^{out}} - \frac{K^{in} d^{in}}{A_0^{in}} = 0$$
$$A_0 \equiv \frac{(K^{out} + K^{in}) A_0^{out} A_0^{in}}{K^{out} A_0^{in} + K^{in} A_0^{out}} , \tag{2.21}$$

where A_0 is defined to be the preferred area of the bilayer neutral surface, and d^{out}, d^{in} indicate the distances of the two monolayer neutral surfaces from that of the bilayer, i.e., $d^{in} + d^{out} = D$. Again, the total bilayer elastic energy should generally contain contributions coming from both area dilation and bending of the bilayer neutral surface. If we characterize the neutral surface of a vesicle by its surface area A, and its local principal curvatures C_1 and C_2, then the closed geometry of the vesicle requires that the area dilation/compression of each monolayer should be

$$\frac{A^{out}}{A_0^{out}} - 1 = \frac{A}{A_0^{out}} - 1 + \frac{d^{out} \oint dA_0 (C_1 + C_2)}{A_0^{out}} + \frac{(d^{out})^2 \oint dA\, (C_1 C_2)}{A_0^{out}}$$
$$\frac{A^{in}}{A_0^{in}} - 1 = \frac{A}{A_0^{in}} - 1 - \frac{d^{in} \oint dA_0 (C_1 + C_2)}{A_0^{in}} + \frac{(d^{in})^2 \oint dA\, (C_1 C_2)}{A_0^{in}} \quad . \tag{2.22}$$

The monolayer mean curvature is simply

$$C_1^{out} + C_2^{out} = (C_1 + C_2) - d^{out}(C_1^2 + C_2^2)$$
$$C_1^{in} + C_2^{in} = (C_1 + C_2) + d^{in}(C_1^2 + C_2^2) \quad . \tag{2.23}$$

Substituting Eqs.(2.22) and (2.23) into the expression (2.20), we finally obtain the elastic-energy functional for the bilayer,

$$\begin{aligned} W &= const. + \frac{(K^{out} + K^{in})A_0}{2}(A/A_0 - 1)^2 \\ &+ \frac{\kappa}{2}\oint dA\left[(C_1 + C_2 - C_0)^2 + O(C_{1,2}^2 D/R)\right] \\ &+ \frac{\bar{\kappa}}{2}\left[\frac{\pi}{AD^2}(\Delta A - \Delta A_0)^2 + O(D/R)\right] \\ &+ \kappa_g \oint dA\left[C_1 C_2 + O(C_{1,2}^2 D/R)\right] \ , \end{aligned} \tag{2.24}$$

where

$$\kappa \equiv \kappa^{out} + \kappa^{in} \ , \tag{2.25}$$

$$\bar{\kappa} \equiv \frac{K^{out}K^{in}A_0}{\pi(K^{out}A_0^{in} + K^{in}A_0^{out})}D^2 \ , \tag{2.26}$$

$$C_0 \equiv \frac{\kappa^{out}C_0^{out} + \kappa^{in}C_0^{in}}{\kappa^{out} + \kappa^{in}} \ , \tag{2.27}$$

$$\kappa_g \equiv \kappa_g^{out} + \kappa_g^{in} + \frac{K^{out}K^{in}(K^{out}A_0^{in} - K^{in}A_0^{out})(A_0^{out} - A_0^{in})}{(K^{out}A_0^{in} + K^{in}A_0^{out})^2}D^2 \ . \tag{2.28}$$

Non-zero values of the bilayer spontaneous curvature C_0 in the free energy (2.24) reflect a spontaneous tendency for the bilayer to curve one way or the other, arising from the asymmetry between the constituent monolayers, as suggested by the simple expression (2.27). If two identical lipid monolayers ($\kappa^{out} = \kappa^{in}$ and $C_0^{out} = -C_0^{in}$)[5] are brought together, the intrinsic monolayer asymmetry cancels completely, and $C_0 = 0$. There are at least two possible mechanisms which can lead to an incomplete cancellation and a nonzero value of C_0. The first invokes a molecular compositional asymmetry between the monolayers,

[5]The negative sign appears because both C_0^{out} and C_0^{in} are now referred to a single inside/outside sign convention, which treats the head/tail asymmetry of the individual monolayers oppositely.

so that $\kappa^{out} \neq \kappa^{in}$, $C_0^{out} \neq C_0^{in}$, simply because of some difference in the chemical structures of lipid molecules on the two sides. The second mechanism allows the two leaves to be identical in composition but focusses on the chemical difference between the adjacent fluid environments inside and outside the vesicle. Such chemical differences, as we have already discussed, ultimately lead to different modifications in steric shapes of molecules–even the same kind of molecules–as they interact with the different fluids inside and outside the vesicle. This effect can also modify the microscopic interactions that are responsible for the macroscopic elasticities.

We comment briefly on the physical meaning of the individual terms in Eq. (2.24) and on their relative contributions to the overall energy. The first term in (2.24) is the elastic energy needed to stretch the bilayer from its preferred area A_0 to its actual overall area A. The factor $(A/A_0 - 1)$ is dimensionless, so the energy scale of this contribution is $KA_0 \sim KR^2$, i.e., larger by a factor $(R/D)^2$ than the other elastic moduli in the problem (Eqs. (2.12) and (2.14), and Eqs. (2.25) and (2.28)). Because of this large energy, deformations involving dilation are ordinarily strongly suppressed. When this is so (as it will usually be, in what follows), it is permissible to drop this term and replace it by the constraint that A be fixed at its preferred value A_0. This is the case, for example, for flaccid vesicles, which can take on many different shapes at fixed V and A, differing in energy from one another only at the energy scale of KD^2 of the bending energies. The only situation in which this does not apply is when large osmotic pressures inflate the vesicle so that its volume becomes greater than $V_0 = (4\pi/3)(A_0/4\pi)^{3/2}$; thus, the preferred membrane area can not accommodate the required volume. Under these conditions, the overall dilation energy enters, the energy scale of fluctuations is KR^2, the large fluctuations characteristic of flaccid vesicles are suppressed, and all vesicle shapes are (strained) spheres.

The local curvature variables occur only in the second and fourth terms. These terms describe the shape-dependent part of the overall bilayer energy. Discussion of these terms for the bilayer model parallels what was said after Eq. (2.9) for the monolayer model. In particular, these terms include all Euclidean invariants in the local curvatures through second order. The integrals in both terms are dimensionless, so the two bending rigidities κ and κ_g

have units of energy. Both these energies are generically of order $K\delta^2$, although , it may be expected that κ_g is appreciably smaller than κ, as explained after Eq. (2.9). The Gaussian term is a topological invariant by virtue of the Gauss-Bonnet theorem ($\oint dAC_1C_2 = 4\pi$ for sphere topology). Thus, in comparing the energies of different shapes with the same topology, this term provides a trivial additive constant and will be omitted henceforth.[6]

Finally, the third term in Eq. (2.24) represents the cost in elastic energy of deformations which force the area difference ΔA of the monolayers to depart from the relaxed value ΔA_0, which it would have if each molecule had exactly its optimal area. This term is *global* or *non-local* in the sense that, because of Eq. (2.2), it involves the integral of the mean curvature over the entire vesicle surface. The factor in the square bracket is dimensionless and of order unity in magnitude for unstrained configurations. The area-difference elasticity $\bar{\kappa}$ (Eq. (2.26)) has a magnitude of order KD^2, which is the same scale as that of κ and κ_g, at least under (typical) conditions, where $D \sim \delta$. The reason for this is demonstrated in the microscopic derivation: This term arises generically from the lateral elastic dilation/compression of the constituent monolayers (Eq. (2.20)). The actual surface areas of the two leaves are constrained by the global geometry, as indicated in Eq. (2.22). Consequently the optimal state, where both monolayers assume their preferred surface areas A_0^{out} and A_0^{in}, will not be reached when the vesicle takes an arbitrary shape. A uniform area dilation/compression of the order of D/R is in general unavoidable for either or both leaves, giving rise to the non-local elastic energy of order of $K(D/R)^2A \sim KD^2$.[7]

The upshot of this discussion is that, for flaccid vesicles of fixed topology, the full elastic energy Eq. (2.24) simplifies to Eq. (2.1). This is the ADE model which we shall study in Chapter 4.

The non-local bending energy (third term in Eq. (2.24); second term in Eq. (2.1)) is

[6] In comparing the energies of topologically different bilayer configurations, this term plays a crucial role [38].

[7] This can be easily illustrated in the following example. We consider a bilayer which is initially flat and its two constituent monolayers have the exactly same optimal area. Imagine the bilayer is then closed to form a vesicle such that the mid-surface of the bilayer is unstretched. The outer leaf must experience an area expansion of $(D/2)\oint dA(C_1+C_2)/A \sim (D/R)$ where the inner one suffers a compression $-(D/2)\oint dA(C_1+C_2)/A$.

absent in the SC and the ΔA models. As shown above by the derivation based on a simple microscopic model, this term is of the same order of magnitude as the local bending elasticity term (Eq. (2.12) and Eq. (2.25)). This expectation is supported by the first direct measurement of non-local bending rigidity by Waugh, Song, Svetina and Žekš. This experiment uses the tether formation technique and yields a value $\bar{\kappa}/\kappa \approx 1.1$ [39]. Thus, there seems no justification for omitting this term in studying equilibrium vesicle shapes.

The non-local elasticity term has a form rather different from that of the local bending energy: It seems to involve some long-range interaction between local curvatures at different points on the surface, whereas it is clear that the interactions we start with are completely local (Eq. (2.16)). It may be, therefore, useful to summarize here the reasons for the apparent non-locality. It comes about because of the homogeneity of the local density distribution of the individual monolayer (Eq. (2.19)), the global constraints imposed by the fixed distance D between the individual monolayers and the closed boundary condition, and the fact that the numbers of lipid molecules in the individual monolayers are separately conserved (see Section 1.1).

The microscopic picture (i.e., the four-layer picture) of a fluid lipid bilayer, on which we have based our derivation and discussion of the ADE model, was first proposed by Svetina, Brumen and Žekš in 1985. However, our presentation in this section is the first "systematic" discussion, which gives a unifying and semi-quantitative (i.e., order-of-magnitude) explanation of the various macroscopic elasticities of a fluid-lipid-bilayer membrane. The energy scale associated with the dilation of the membrane surface, KR^2, and those associated with the bending, KD^2, are all determined by the elementary elastic constant, K, the area-compressibility modulus, and by the two fundamental length scales in the system, R, the macroscopic scale, and D, the microscopic thickness of the membrane.

We complete the discussion of the ADE model Eq. (2.1) by summarizing all the constraints in this model. The ADE model takes as constraints fixed values of the surface area and the enclosed volume. There is no constraint on the *geometrical* area difference ΔA between the two leaves of the bilayer. The *optimal* area difference ΔA_0 depends on the history of the formation of the vesicle and is a shape-independent constant, as long as

temperature and other macroscopic conditions in the adjacent fluids are kept fixed. The spontaneous curvature C_0 is also assumed constant over the whole vesicle surface and is independent of the specific vesicle shape profile S.[8]

2.2 Two canonical continuum models–special cases of the ADE model

The discussion of the ADE model in the preceding section enables us to introduce the two canonical models, the spontaneous-curvature model and the area-difference model, as special cases of the ADE model. In this process, both the physics they capture and the intrinsic limitations they have can be illustrated in a transparent way.

The SC model describes a closed fluid lipid-bilayer vesicle by the energy functional

$$E[S] \equiv \frac{\kappa}{2} \oint dA[C_1(\mathbf{r}) + C_2(\mathbf{r}) - C_0]^2 \,, \tag{2.29}$$

supplemented by constraints of fixed surface area and volume.

The SC model was first introduced by Canham [6] and subsequently has been reformulated by Helfrich [7] to take into account the effect of bilayer asymmetry via the spontaneous curvature. Helfrich and Deuling [8] exploited the model further: They found a catalogue of vesicle shapes, including prolate ellipse and oblate ellipse, and showed that, at given vesicle volume and area, prolate and oblate ellipses have different bending energies. Many others have also made their particular contributions [32, 40, 41, 42]. For example, Peterson [42] performed a stability analysis of vesicle shapes which indicated that the axisymmetric shapes are frequently but not always stable with respect to non-axisymmetric perturbations.

The energy functional (2.29) contains only a local bending-energy term. As we have learned from the derivation presented in the previous Section (see Eqs. (2.5) and (2.9)), this energy originates from the bending deformation of two surfaces which are coupled locally. The SC model therefore implies that the two monolayers forming the bilayer are locally constrained, both dynamically and statically, thus behaving effectively as one, and

[8]This may not always be true. Indeed, when the vesicle bilayer is composed of a mixture of lipids and there is surface domain formation, one can expect this parameter to vary locally.

it simply does not see the bilayer structure of the membrane. However, it has been shown that, for bilayer vesicles composed of glycerophospholipids, the two constituent monolayers are not laterally constrained to each other [43]. The description given by the SC model is therefore not fully applicable to these systems.

The ΔA model (originally called the bilayer-couple model) takes a different perspective. Like the SC model, it is described by a bending energy depending upon the local curvatures of the vesicle surface,

$$G[S] \equiv \frac{\kappa}{2} \oint dA[C_1(\mathbf{r}) + C_2(\mathbf{r})]^2 \,. \tag{2.30}$$

The constraints in this model include the fixed surface-area and volume constraints. In addition, there is a third constraint on the monolayer-area difference ΔA.

This model was put forward by Svetina and Žekš, based on the bilayer-couple hypothesis: The two leaflets of a closed bilayer membrane may respond differently to various perturbations while remaining in contact [30, 44]. Subsequent investigations of this model by Svetina and Žekš [25, 26, 31] have indicated a number of interesting properties of vesicle shapes, such as the occurrence of the symmetry-breaking instability associated with discocyte–stomatocyte transformation and geometrical limitations of possible shapes. The full picture of shapes and shape transformations has been given recently by Seifert, Berndl and Lipowsky [27].

The ΔA model does see the two constituent monolayers as independent entities. It assumes that the optimal state, where both of the two leaves have their preferred surface areas, is strictly enforced upon the vesicle, thereby considering each lipid monolayer absolutely unstretchable or incompressible. This leads to the additional global constraint, $\Delta A \equiv \Delta A_0$. Under the assumption that the bilayer thickness D is a constant over the entire surface, this in turn leads to a constraint on the integrated mean curvature, as follows from Eq. (2.2). The spontaneous curvature C_0 is set to zero, since it is coupled to the constrained ΔA, thus merely shifting the curvature energy by a constant, and is, therefore, irrelevant to equilibrium shapes. However, the assumption that ΔA is constrained is not a very realistic one. Indeed, as we have pointed out, a deviation of each monolayer area from

its optimal value of the order of D/R should be possible and will make contributions to the total elastic free energy which are comparable to the local bending energy. The final value of ΔA should be determined as an outcome of minimizing the total free energy instead of being fixed *a priori*.

Both of these curvature-energy models can be regarded as the special cases of the ADE model, wherein the non-local bending effect is looked at from two extreme points of view. The SC model corresponds to the case where the non-local bending constant $\bar{\kappa}$ is taken to be zero, whereas the ΔA model is recovered as the limiting case when $\bar{\kappa} \to \infty$. Furthermore, neither the SC model nor the ΔA model gives a unifying explanation of all the experimental findings. A more detailed discussion of this issue is deferred to Chapter 4.

2.3 Concluding remarks

We conclude this chapter with some general remarks on the continuum models. First, all three of the curvature-energy models we have considered in this chapter have neglected short-distance effects (i.e., effects at distances comparable to the thickness of the lipid bilayer), where van der Waals forces may be an important source for intravesicular interaction. Indeed, Bruinsma [45] has recently shown how van der Waals interactions may stabilize tubular structures and/or cause dynamical instability towards bead-like structures, which, under appropriate circumstances, might resemble our equilibrium multiplets (see Chapter 3). Secondly, these models assume that the bilayer under consideration has a homogeneous surface structure. Important modifications will have to be made when one investigates the equilibrium shapes of vesicles consisting of a lipid mixture which is capable of phase separating to form intramembrane domains [46].

Chapter 3

Equilibrium Shapes and Shape Transitions in the Spontaneous-Curvature Model

3.1 Introductory remarks

In this chapter, we describe the theory of equilibrium shapes and shape transitions for fluid lipid-bilayer vesicles within the context of the SC model. In particular, we shall concentrate on the equilibrium shapes and shape transitions involving "budding" and "vesiculation" processes.

"Budding" and "vesiculation" refer to processes in which a single, more or less spherical vesicle (or cell in the biological context), when subjected to change in some external parameter or stimulus, undergoes a shape transformation to produce one or more globular appendages linked to the parent body through narrow necks or tubes. When the neck has a non-zero diameter, we shall call the process "budding". To distinguish the limiting case where the neck diameter goes to zero[1], we shall refer to this as "vesiculation". Both these processes affect the shape but not the topology. If subsequently the appendage breaks off from the parent body (fission), only then does the topology change.

[1] "Zero" neck diameter is a well-defined limit in the continuum description. In real life, it corresponds to a short-distance cut-off, at which the continuum description may fail, and van der Waals interactions may become important.

Budding of small vesicles from large membrane surfaces was originally observed in biological cells. It is a frequent event in cells and is usually driven by complex chemical stimuli [47]. However, budding has also been found to occur in non-biological systems, such as artificial laboratory vesicles, as shown by the controlled experiments reviewed in Chapter 1. It is the theme of this chapter that budding and vesiculation in artificial vesicles can be successfully described within the context of continuum models.

The equilibrium vesicle shape is that which minimizes the SC bending energy (2.29) at given A and V, i.e., it is the mechanically stable shape of lowest energy. For many experimental systems (e.g., fluid bilayers made of SOPC and DMPC), the bending rigidity κ is appreciably larger than the thermal energy at room temperature,[2] so thermal fluctuations are small, and this zero-temperature ($T = 0$) approximation is valid. Under these conditions, the energy scale κ drops out and only geometry is left. There remain three length scales, R_A, R_V, R_0, defined by $A \equiv 4\pi R_A^2$, $V \equiv 4\pi R_V^3/3$ (note that $R_A \geq R_V$), and $C_0 \equiv 2/R_0$. We shall assume that $C_0 \neq 0$.[3] There is then no loss of generality[4] in taking $C_0 > 0$, so spheres of radius R_0 (which we shall refer to as "Helfrich spheres") cost no energy, and we are free to set the scale of length by taking $C_0 = 1$ (i.e., $R_0 = 2$).

The minimization of the mechanical bending energy expressed by Eq. (2.29) at fixed A and V leads to a catalog of shapes, which was first compiled by Deuling and Helfrich [8] and has subsequently been extended and refined by other authors [40, 41, 42]. Many (but not all [25, 26, 27, 31]) authors have restricted their attention to the regime $R_A, R_V \sim 1$, i.e., vesicle dimensions comparable to R_0. A variety of shapes and shape transitions do, indeed, show up in this region, including discocytes, stomatocytes, torocytes (mostly corresponding to negative values of C_0), etc.; however, budding and vesiculation are absent.

The purpose of our investigation of the spontaneous-curvature model was, therefore, twofold. First, we wished to explore the regime $R_A \sim R_V \gg 1$, of large and more or less full vesicles. In this we were motivated by the early experimental observations of Evans

[2] Typical values for κ quoted [22] for PC-lecithins are $0.4 - 2.0 \times 10^{-12} erg$ or $(10 - 20) k_B T_{room}$, in agreement with the rough estimate given in the microscopic derivation presented in the preceding chapter.

[3] The analysis we shall perform is perfectly applicable to the case where $C_0 = 0$; however, no budding and vesiculation occurs in that case.

[4] $C_0 < 0$ is achieved in this convention by closing the vesicle "inside out", i.e., below $z = 0$ in Fig. 3.1.

and Rawicz [18], Sackmann, Duwe, and Engelhardt [48] and others [23] on budding of giant artificial vesicles. Our second focus was to construct for the above regime a real "phase diagram," showing the systematics of shapes and shape changes, rather than just a catalog of shapes: The systematics is crucial to testing the different theoretical models, since, as it turns out, all models give same catalog of shapes. By simply looking at equilibrium shapes, one is not able to distinguish one model from the other. It is the fundamental outcome of our study that budding and vesiculation can and do both occur as adiabatic processes (i.e., processes involving only equilibrium states) in the SC model–the simplest model for fluid vesicles.

The physical origin of budding and vesiculation in this continuum curvature model is easy to understand. When $R_A = R_V$, the only possible vesicle shape is a sphere. As area increases via, e.g., thermal expansion (or, equivalently, as volume decreases through osmotic deflation), an excess area A_Δ becomes available and other shapes can occur. If we write $A = 4\pi R_V^2(1+\Delta)$ in terms of the fractional excess area Δ, then $A_\Delta = 4\pi R_V^2\Delta$. As soon as A_Δ becomes comparable to $4\pi R_0^2$ (16π in our units), it is favorable for the vesicle to shed the excess area in the form of a bud with a radius of the order of R_0, which costs very little energy (we emphasize again that no energy is required when the radius is exactly R_0).[5] As the area increases further, the shedding process proceeds in a more or less periodic manner, until the volume of the parent body becomes, itself, comparable to that of the Helfrich sphere. The full sequence of shape transitions is quite complicated, and we will find that it involves a nice interplay of some shape configurations with small necks (buds) and others in which the neck radii shrink to zero (vesiculation).

This mechanism for budding and vesiculation appears to have been originally proposed by Luke and Kaplan [49]; however, no systematic calculations were done prior to our investigation. Seifert and co-workers independently studied these phenomena from a point of view similar to ours [27]. They have carried out a comprehensive study of the low-volume regime

[5]One might expect that the formation of a narrow neck, at which the two local principal curvatures assume large values, requires a large curvature energy. It turns out that, an appropriate neck costs very little energy indeed, no energy as the neck radius $a \to 0$, as we shall discuss in detail in Subsection 3.2.3 and Appendix C.

(i.e., $V \sim 4\pi R_0^2$). Budding and vesiculation do occur in their studies, but the sequences which we shall find at high volume are absent.

We have argued in Chapter 2 that the SC model is not fully adequate for describing lipid-bilayer vesicle shapes. What, then, is the rationale for studying this model in detail? The answer is several-fold:

(a) The SC model was the first and for some time the only model of fluid vesicle shapes. At the time when this work was done, it was still regarded as a credible model of equilibrium vesicle shapes. Indeed, it is at least partially a consequence of our work on the shape-change systematics that it was possible to see that the SC model did not–apparently–adequately describe what was seen in the lab for giant artificially prepared phospholipid vesicles.

(b) The SC model does correctly capture some of the physics of fluid lipid bilayers, even though it does not capture all of it. It does recognize the central importance of the local bending elasticity. It is of theoretical interest, at least, to understand what implications this crucial term has–acting by itself–for vesicle shapes and shape transformations.

(c) There exist two-headed lipids, called bola lipids [50]. In a lipid bilayer containing a significant number of bola lipids, the two leaves may be expected to be "spot-welded" together, thus preventing free sliding of the two leaves relative to one another. The SC model would be a credible model for such systems. Indeed, there is at least preliminary evidence that the kind of multiple vesiculation predicted by the SC model is especially common when bola lipids are present [50].

(d) Most of the ideas and concepts for the description of equilibrium vesicle shapes and shape transitions developed within the framework of the SC model are fundamentally and universally important to this research field.

(e) The systematic understanding of the SC model will shed light on our investigation of the more general ADE model, as it is one of the two limiting cases of the ADE model. In fact, it will become clear in the next chapter that a great technical advantage can be gained from studying such limiting cases.

3.2 General formulation of the problem and systematics of the analysis

3.2.1 Constrained minimization as a variational problem

Following Deuling and Helfrich [8], we formulate the constrained minimization of the bending energy (2.29) as a variational problem by introducing the variational functional,

$$\Phi[\Sigma, P; S] = E[S] + \Sigma A[S] - PV[S] \,, \tag{3.1}$$

which we shall refer to as a free energy, in analogy with thermodynamics. Σ and P are Lagrange multipliers used to enforce the surface-area and volume constraints, respectively. Σ has the dimension of surface tension, while P has the dimension of pressure. In what follows, we shall choose units such that κ and C_0 are equal to unity. The Lagrange multipliers σ and p are thus dimensionless, defined as,

$$\sigma \equiv \Sigma/\kappa C_0^2 \,, \quad p \equiv P/\kappa C_0^3 \,. \tag{3.2}$$

We shall restrict our attention, for technical reasons, to shapes with axisymmetry. Of course, it is not always true that a mechanically stable axisymmetric shape will be stable against local non-axisymmetric perturbations, even if it is stable with respect to axisymmetric perturbations. This restriction may be justified, in principle, only if a careful local analysis of non-axisymmetric perturbations about an axisymmetric shape proves its stability (Refs. [51, 52, 53]). On the other hand, this restriction is not too limiting in practice: Most of the vesicle shapes observed in experiments (Fig. 1.5) exhibit axisymmetry in the regime of parameters which we shall explore. Theoretical work is consistent with this expectation [51].

Parametrization of an axisymmetric shape is illustrated in Fig. 3.1. z and r measure distances along and perpendicular to the symmetry axis, respectively. The function $z(r)$ then determines a shape. Let Θ be the angle between the local tangent and the symmetry axis, and $C_m(r)$, $C_p(r)$ be the principal curvatures, defined positive when the center of curvature lies along the direction of the interior normal $\hat{n}$. It is the consequence of the

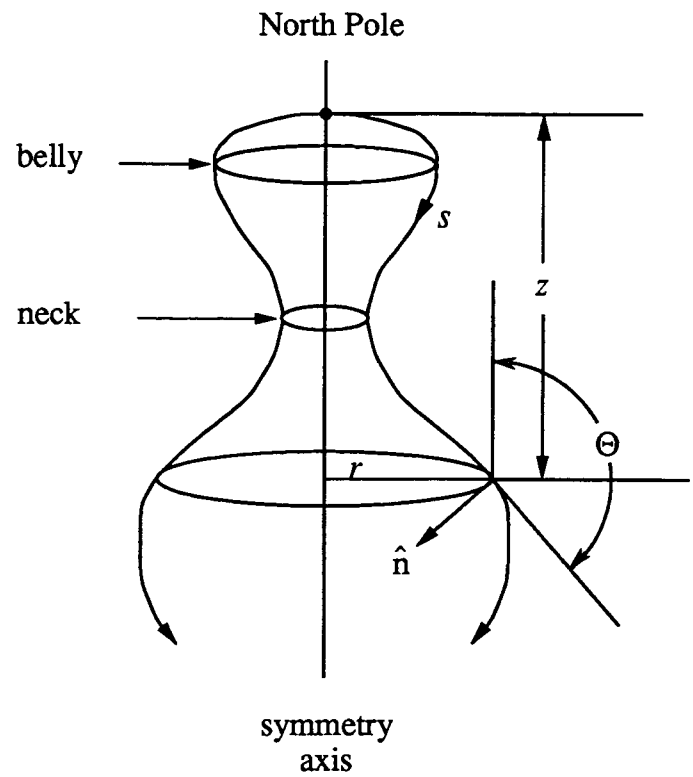


Figure 3.1: Parametrization of an axisymmetric vesicle shape. Either the distance from the axis, r, or the arclength, s, may be used as the variable parametrizing the shape. The local principal curvatures C_m and C_p are defined positive when the corresponding center of curvature lies along the line of the interior normal $\hat{n}$.

axisymmetry that

$$C_m = -\frac{d(\cos\Theta)}{dr}\,, \quad C_p = -\frac{\cos\Theta}{r}\,, \tag{3.3}$$

where $\Theta(r)$ depends on $z(r)$,

$$\cot\Theta = \frac{dz}{dr}\,. \tag{3.4}$$

Making the free energy (3.1) stationary subject to the condition of axisymmetry leads to the following set of Euler-Lagrange equations, originally derived by Deuling and Helfrich

[8] (the "'Helfrich representation"),

$$\begin{aligned}\frac{dC_m}{dr} &= \frac{r}{2[1-(rC_p)^2]}\left\{C_p\left[(C_p-1)^2-C_m^2\right]+2\sigma C_p-p\right\}-\frac{C_m-C_p}{r}\,,\\ \frac{dC_p}{dr} &= \frac{C_m-C_p}{r}\,.\end{aligned} \tag{3.5}$$

This set of Euler equations is convenient for analytical treatment of the problem, as will be demonstrated in our discussion of the nearly vesiculated shapes (Appendix C); however, it has the deficiency that points with infinite derivative dz/dr ("belly" points) are singular points. Near these points, numerical integration of the Euler equations cannot be dealt with efficiently.

This problem was solved by Peterson [42]: Using the arclength s as the parametrization variable, he first composed the following variational energy,

$$\begin{aligned}\Psi[\sigma,p,b;S]/\pi &\equiv \Phi[\sigma,p;S]/\pi+\int_0^{s_0} ds\, b(s)\left[\frac{dr}{ds}-\sin\Theta\right]\\ &= \int_0^{s_0} ds\left\{r\left[\frac{d\Theta}{ds}-\frac{\cos\Theta}{r}-1\right]^2+2\,\sigma\, r+p\,r^2\cos\Theta+b(s)\left[\frac{dr}{ds}-\sin\Theta\right]\right\}\quad,\end{aligned} \tag{3.6}$$

following from the geometrical relationships,

$$\begin{aligned}\frac{dr}{ds}&=\sin\Theta\,,\quad \frac{dz}{ds}=\cos\Theta\,,\\ C_m&=\frac{d\Theta}{ds}\,,\quad C_p=-\frac{\cos\Theta}{r}\,.\end{aligned} \tag{3.7}$$

The last term in this variational energy was introduced specifically to deal with the local constraint, $dr/ds = \sin\Theta$, imposed by the geometry on the functions $r(s)$ and $\Theta(s)$. The function $b(s)$ may be considered a "Lagrange-multiplier function" which enforces this constraint.

Making the variational energy (3.6) stationary against arbitrary (axisymmetric) shape changes leads to

$$\frac{dz}{ds} \;=\; \cos\Theta\,, \tag{3.8}$$

$$\frac{dr}{ds} = \sin\Theta\,, \tag{3.9}$$

$$\frac{d\Theta}{ds} = C_m\,, \tag{3.10}$$

$$\frac{dC_m}{ds} = -\frac{C_m\sin\Theta}{r} - \frac{\cos\Theta\sin\Theta}{r^2} - \frac{p}{2}\,r\sin\Theta - \frac{1}{2}\frac{b\cos\Theta}{r}\,, \tag{3.11}$$

$$\frac{db}{ds} = 2p\,r\cos\Theta + 2\sigma + (C_m - 1)^2 - \frac{\cos^2\Theta}{r^2}\,, \tag{3.12}$$

along with a conserved "Hamiltonian function"

$$\mathcal{H}(s) \equiv r\left[\left(\frac{d\Theta}{ds}\right)^2 - \left(-\frac{\cos\Theta}{r} - 1\right)^2\right] - 2\sigma r - p\,r^2\cos\Theta + b\sin\Theta\,, \tag{3.13}$$

which is identically zero along the shape contour.[6] The boundary conditions for Θ and r at both the north and south poles are obvious,

$$\begin{aligned} &\Theta(0) = \pi/2\,, \quad \Theta(s_0) = 3\pi/2\,, \\ &r(0) = 0\,, \qquad r(s_0) = 0\,. \end{aligned} \tag{3.14}$$

Note that $z(0)$ can be chosen to be 0 as a convention, while $z(s_0)$ depends on the particular shape calculated. It follows from the conservation of $\mathcal{H}$ that the boundary conditions for $b(s)$ must be

$$b(0) = 0\,, \quad b(s_0) = 0\,. \tag{3.15}$$

We shall henceforth refer to this set of Euler equations as the "Peterson representation". The belly points are no longer the singular points of this set of Euler equations. We therefore performed all our numerical calculations of the shapes using the Euler equations in this representation.

In Appendix A, we give detailed derivations of the Euler-Lagrange equations in both representations and we show their equivalence.

[6] There is nothing mysterious about this Hamiltonian function: It comes about as a conserved quantity, simply because the integrand in the variational energy (3.6) has no explicit dependence on the variable s.

3.2.2 Euler shapes and the phase diagram

Smooth solutions to the Euler-Lagrange equations, Eq. (3.5) or Eqs. (3.8)–(3.12), define what we shall call the Euler shapes, i.e., the axisymmetric shapes that are in mechanical equilibrium. These shapes are found by integrating the equations for $C_m(s)$ and $C_p(s)$ from $s = 0$ (the "north pole") to $s = s_0$ (the "south pole"). Boundary conditions at the north pole are $r = 0$ and $C_m = C_p$. Suppose that the integration starts at $s = 0$ with initial curvature $\alpha_0 = C_m(0) = C_p(0)$. For a general value of α_0, the figure will not close and there is no Euler shape. For each value of σ and p there are only a discrete set of initial curvatures $\{\alpha_0^{(n)}(\sigma, p)\}$ for which the shape equations close smoothly at the south pole. We shall describe in Appendix B systematic algorithms which lead in principle to a complete set of solutions corresponding to the $\{\alpha_0^{(n)}\}$. These solutions, which are the stationary points of the free energy (3.1), are local minima or saddle points in the space of all configurations. The local minima correspond to shapes which are locally stable; the saddle points give the activation energies for transitions between different minima. Whether a shape is a minimum or a saddle point can be investigated by performing a local stability analysis. In this work we have not attempted to perform such an analysis in general;[7] however, we shall address the question of stability near the spherical limit in Section 3.3 below.

The set of Euler shapes, which we denote $S_n(\sigma, p)$, have corresponding areas $A_n(\sigma, p) \equiv A[S_n(\sigma, p)]$, volumes $V_n(\sigma, p) \equiv V[S_n(\sigma, p)]$, and bending energies $E_n(\sigma, p) \equiv E[S_n(\sigma, p)]$. The stationary free energy is then given by

$$\Phi_n(\sigma, p) \equiv \Phi[\sigma, p; S_n(\sigma, p)] = E_n(\sigma, p) + \sigma A_n(\sigma, p) - pV_n(\sigma, p) \, . \tag{3.16}$$

Because $S_n(\sigma, p)$ makes $\Phi[\sigma, p; S]$ stationary, it is easy to show that

$$\frac{\partial \Phi_n(\sigma, p)}{\partial \sigma} = A_n \, , \quad \frac{\partial \Phi_n(\sigma, p)}{\partial p} = -V_n \, . \tag{3.17}$$

To do this, note that the variation of $\Phi[\sigma, p; S]$ must vanish,

[7] Note that the Euclidean invariance of the bending energy functional and the constraints guarantees that shapes which satisfy the axial Euler equation are stationary with respect to local non-axial perturbations. It is not guaranteed, however, that such stationary shapes are local minima, and, indeed, it is not always true. This point has been explored by Peterson, Milner and Safran, Ou-Yang and Helfrich in special cases (Refs. [51, 52, 53, 54]).

$$
\begin{aligned}
\delta\Phi &= \Phi[\sigma,p;S_n(\sigma+\Delta\sigma,p+\Delta p)]-\Phi[\sigma,p;S_n(\sigma,p)] \\
&= E_n(\sigma+\Delta\sigma,p+\Delta p)+\sigma A_n(\sigma+\Delta\sigma,p+\Delta p)-pV_n(\sigma+\Delta\sigma,p+\Delta p) \\
&\quad -[E_n(\sigma,p)+\sigma A_n(\sigma,p)-pV_n(\sigma,p)] \\
&\equiv \frac{\delta\Phi}{\delta\sigma}\Delta\sigma+\frac{\delta\Phi}{\delta p}\Delta p=0\,. \qquad (3.18)
\end{aligned}
$$

This leads to

$$
\begin{aligned}
\frac{\delta\Phi}{\delta\sigma} &\equiv \frac{\partial E_n}{\partial\sigma}+\sigma\frac{\partial A_n}{\partial\sigma}-p\frac{\partial V_n}{\partial\sigma}=0 \\
\frac{\delta\Phi}{\delta p} &\equiv \frac{\partial E_n}{\partial p}+\sigma\frac{\partial A_n}{\partial p}-p\frac{\partial V_n}{\partial p}=0\,. \qquad (3.19)
\end{aligned}
$$

On the other hand, we may calculate the differential change of Φ_n as,

$$
\begin{aligned}
\Delta\Phi_n &\equiv \Phi[\sigma+\Delta\sigma,p+\Delta p;S_n(\sigma+\Delta\sigma,p+\Delta p)]-\Phi[\sigma,p;S_n(\sigma,p)] \\
&= \left[\frac{\delta\Phi}{\delta\sigma}+A_n(\sigma,p)\right]\Delta\sigma+\left[\frac{\delta\Phi}{\delta p}-V_n(\sigma,p)\right]\Delta p \\
&\equiv \frac{\partial\Phi_n}{\partial\sigma}\Delta\sigma+\frac{\partial\Phi_n}{\partial p}\Delta p\,. \qquad (3.20)
\end{aligned}
$$

Combining this with Eq. (3.19), we arrive at the result given in (3.17). There are typically many solutions to the Euler equations at given σ and p. The evolution of this set of solutions as σ and p vary leads to a family of sheets described by $\{\Phi_n(\sigma,p)\}$[8] over the (σ,p) plane. On a single sheet, the dependence $\Phi_n(\sigma,p)$ is generally analytic; however, the sheets can cross, they can fold over and connect to each other, and they can merge into or bifurcate from one another.

If the free energy $\Phi_n(\sigma,p)$ is not globally convex, there may be several solutions to the inversion $A_n(\sigma,p)=A$, $V_n(\sigma,p)=V$ for each sheet n at given A, V. Let $\sigma_m(A,V),p_m(A,V)$

[8]It may be asked at this point why one does not form directly the energy surfaces $E_n(\sigma,p)$ rather than the free-energy surfaces. The answer is that the Legendre-transform structure makes Φ the natural free energy. Relations like Eq. (3.17) and (3.22) mean that local second-order structure is normally preserved (but see the next footnote for an exception).

and $S_m(A,V)$ label the complete set of such solutions (i.e., over both n and the different solutions for each fixed n). $(\sigma_m(A,V), p_m(A,V))$ consists of a discrete set of points. The corresponding family of energy surfaces,

$$E_m(A,V) = \Phi\left[\sigma_m, p_m; S_m(A,V)\right] - \sigma_m A + p_m V \ , \tag{3.21}$$

is simply a Legendre transform of the stationary free energy Φ_n and has, also, a complex structure. Local bifurcation structure is normally, but not always[9], preserved in the change of variables. It is a consequence of the Legendre-transform structure that

$$\frac{\partial E_m(A,V)}{\partial A} = -\sigma_m \ , \quad \frac{\partial E_m(A,V)}{\partial V} = p_m \ , \tag{3.22}$$

which is true as a local statement for any sheet m. Although Eqs. (3.17) and (3.22) look very "thermodynamic," it must be kept in mind that the functions $\Phi_n(\sigma,p)$ and $E_m(A,V)$ do not in general satisfy convexity conditions even locally, so second derivatives such as $\partial^2 E_m/\partial V^2$ and $\partial^2 E_m/\partial A^2$ may have either sign.

Once all the energy sheets are known (and we shall see in the next section that the Euler-shape sheets are only a subset), the solution to the constrained minimization problem originally posed is

$$E_{\rm eq}(A,V) = \min_S E[S]|_{A,V} = \min_m E_m(A,V) \ , \tag{3.23}$$

i.e., the solution always chooses the lowest-energy sheet at each point (A,V).[10] In regions of the (A,V) plane where the solution remains on a single sheet E_m, the energy $E(A,V)$ is analytic. Nonanalyticities arise when the global minimum $E_{\rm eq}(A,V)$ passes from one sheet onto another. This occurs where sheets cross or where they merge, bifurcate or terminate in some other way. The energy is continuous over the accessible region of the (A,V) space;[11] however, as we shall see in some detail, it is not generally convex and its derivatives may be

[9]The exception is the sphere sheets $\Phi_S(\sigma,p)$ which transforms into a line in $E(A,V)$, thus changing the associated bifurcation structure significantly. We shall discuss this in some detail in Section 3.3. This apparent anomaly is actually a simple consequence of the Euclidean invariance of the problem.

[10]Note that the lowest sheet $E_m(A,V)$ does not necessarily correspond to the lowest sheet $\Phi_n(\sigma,p)$ at the corresponding values $\sigma_m(A,V)$ and $p_m(A,V)$.

[11]The continuity of $E(A,V)$ can presumably be proved rigorously by arguing that a shape S which has energy E at area A and volume V can always be slightly modified in such a way as to acquire area $A+\delta A$ and volume $V+\delta V$ with an energy change δE which goes to zero as δA and δV go to zero.

discontinuous. Borrowing from thermodynamics the terminology of phase transitions , we call the regions of analyticity "single-phase" regions and refer to the loci of nonanalyticity as "phase boundaries". When $\partial E/\partial A$ and/or $\partial E/\partial V$ changes discontinuously, we refer to the transition as "first order" (generically, this is the situation when sheets cross); otherwise, it is "second order" (as occurs at bifurcations). The loci of nonanalyticity in the (A, V) plane constitute the "(A, V) phase diagram". The corresponding loci in the (σ, p) plane constitute the "(σ, p) phase diagram", which nicely provides a Legendre conjugate of the (A, V) phase diagram. Note that crossing a first-order transition corresponds to a discontinuous jump in (σ, p) and may occur at a point where the surface $\Phi_n(\sigma, p)$ is smooth.

3.2.3 Boundary minima, coexistence, and multiplets

It might seem that all minimum-energy shapes correspond to solutions of the Euler equations (3.8)–(3.12). To understand why this is not so, we consider an arbitrary smooth function f of a single scalar variable x in the interval $[b, c]$. To find the minimum of f, it is true that we must search the points where $df/dx = 0$ (analogous to the Euler equation); but, we must also compare with the boundary values $f(b)$ and $f(c)$. These "boundary" minima correspond in the vesicle shape problem to configurations which consist of two or more Euler shapes which touch each other tangentially and possess the required total area and volume collectively (see Fig. 3.2(a)). Technically, such a shape should be thought of as the limit of a single connected (topologically spherical) shape with one or more necks which approach zero diameter.[12] At first sight it might appear that the large curvatures which occur near the necks would always lead to large energies in the bending energy (2.29) and would place such configurations very far from any energy minimum. (Indeed, this is true for a $D = 1$ vesicle in two space dimensions, and for this system vesiculation is absent.) The key point is that, although the absolute values of the two curvatures diverge as the neck diameter a tends to zero, the signs are opposite. It turns out that, by an appropriate choice of neck shape, the leading divergences of the two principal curvatures cancel, and the changes in both the free energy and the bending energy resulting from the opening of a

[12] If fission occurs, then the omitted Gaussian curvature terms would have to be put in.

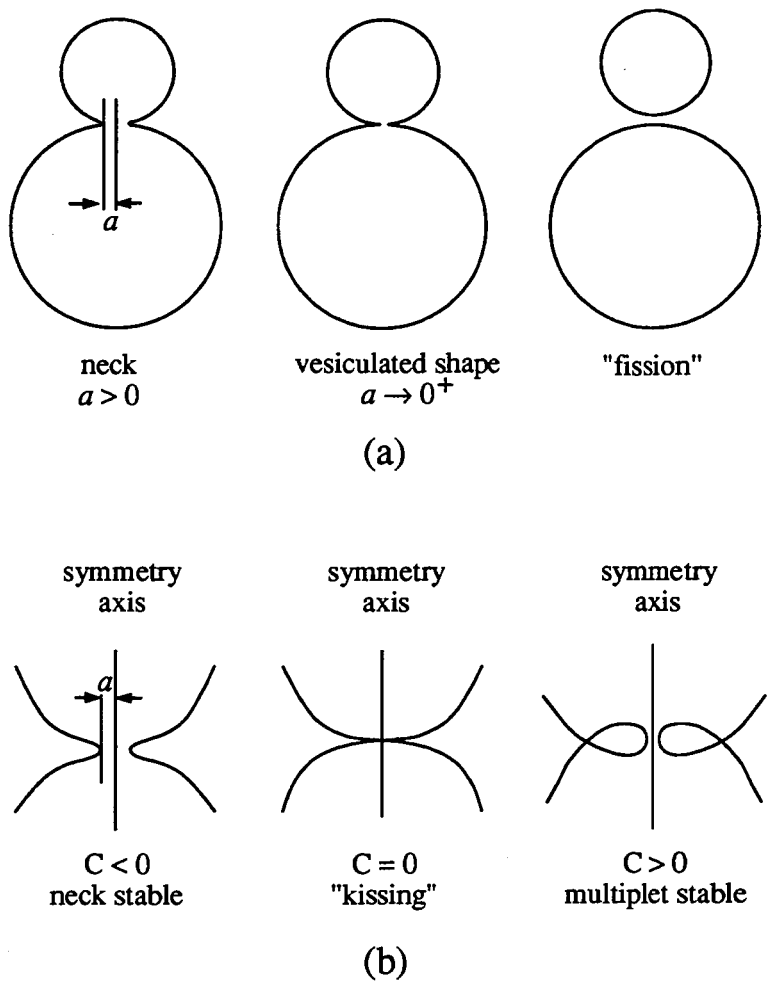


Figure 3.2: Vesiculated shapes versus narrow-necked shapes, and the kissing condition. (a) Vesiculated shapes consist of two or more bodies connected via infinitesimal necks. In situations where these shapes are locally stable (e.g., when $C > 0$ for sphere multiplets), increasing the neck radius a from zero raises the energy. "Fission" changes the topology and is not considered in this thesis. (b) Generic neck geometry of Euler shapes near a kissing boundary ($C = 0$). The criterion $[1 - (1/R_A + 1/R_B)] < 0$ for neck stability applies when the necks join nearly spherical shapes.

neck of size a scale as Ca to lowest order in a (see Appendix C). As long as the coefficient C is positive, the limiting configuration is locally an energy minimum and can compete with the lowest-energy Euler shape of the same area and volume. We shall call these limiting configurations, in which two or more Euler shapes touch tangentially and communicate via an infinitesimal neck, "multiplets".

We give a more detailed account of our treatment of the neck region in Appendix C and simply quote here the three results that are relevant to our present discussion. We have shown (for axisymmetric narrow-necked shapes that approach multiplets consisting of *spherical* segments) the following:

(a) In the limit $a \to 0$, the energy (both the free energy and the mechanical bending energy) of the neck region vanishes.

(b) The coefficient C is proportional to $(1 - 1/R_A - 1/R_B)$ with a positive prefactor, where R_A and R_B are the radii of the corresponding osculating spheres.[13] The multiplet is locally stable when this bracket is positive; however, when this bracket is negative and small in magnitude, a narrow-necked shape which is very close in shape to a spherical multiplet exists as a local minimum.

(c) The equilibrium neck size can be driven to zero by varying parameters such as A and V (or σ and p). At this boundary, the coefficient C becomes equal to zero, so

$$\frac{1}{R_A} + \frac{1}{R_B} = 1 \,, \tag{3.24}$$

which we call the "the kissing condition."

This equation defines lines in the (σ, p) plane (see, e.g., Fig. 3.3).[14] Near such a line there are two sheets of similar stationary-energy shapes, one necked and one vesiculated (Fig. 3.2(b)). The two sheets meet tangentially at the kissing boundary. On one side of the kissing line, $C < 0$ and the necked shapes are locally stable and have lower energy

[13] This corresponds in unnormalized units to $(2/R_0 - 1/R_A - 1/R_B)$, where R_0 is the spontaneous curvature.

[14] Actually, as we show in Appendix C, this kissing condition is of a very local character, in the sense that it only depends upon the local curvatures at the osculating point, not on the global shape of the osculating objects. Hence, we expect a bifurcation structure involving both narrow-necked shapes and vesiculated multiplets whenever the more general condition $C_A + C_B = 1$ is satisfied by two arbitrary axisymmetric Euler objects touching each other tangentially at the symmetry axis. C_A and C_B are local curvatures of the two objects at the osculating point.

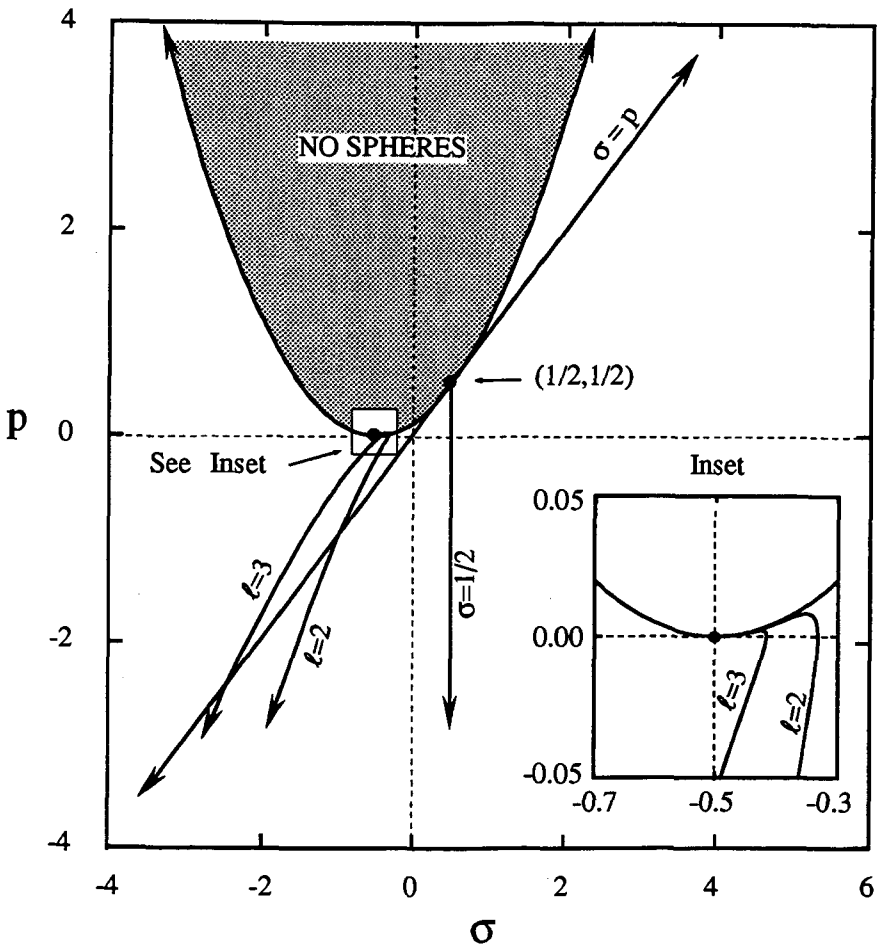


Figure 3.3: The (σ, p) diagram showing special features associated with spheres. Euler spheres do not exist in the region above the parabola, (3.30). Below, there are two distinct spheres with radii given by (3.29). On the parabola, the two radii become equal. The lines labeled by $l = 2$ and $l = 3$ mark the bifurcation boundaries, defined by (3.38), at which small perturbation dominated by a particular l spherical harmonic develops smoothly from spheres. The inset shows how all these bifurcation lines emerge from the point (-1/2,0). For $l = 2$, prolate ellipses lie to the right of the bifurcation line. $\sigma = 1/2$ is the kissing boundary for unequal spheres. $\sigma = p$ is the locus of Helfrich spheres, at which the kissing condition is automatically satisfied.

than the multiplets; while, on the other side, $C > 0$ and the vesiculated shapes are local boundary minima. In this region the neck shape predicted by the Euler-Lagrange equations is self-intersecting and, therefore, unacceptable for a real vesicle.

We now discuss the thermodynamics of the multiplet configurations. A multiplet consisting of M (≥ 2) smooth shapes may be taken to satisfy the following conditions:

$$E(A,V) = \sum_{i=1}^{M} E_i(A_i, V_i) \,, \tag{3.25}$$

$$A = \sum_{i=1}^{M} A_i \,, \quad V = \sum_{i=1}^{M} V_i \,, \tag{3.26}$$

where $i = 1, ..., M$ labels the members of the multiplet. Notice that, in omitting any energy contribution due to the necks, we have made use of result (a) above. In looking for an energy minimum, there is no loss of generality in choosing multiplet members to be Euler shapes, since the lowest-energy *smooth* shape of given area and volume is an Euler shape. It remains to determine how to minimize the energy (3.25) by redistributing the area and volume among all the members of the multiplet subject to the constraints (3.26). Incorporating the constraints via Lagrange multipliers and using Eq. (3.22) for the individual multiplet members yields

$$\sigma_i \equiv -\frac{\partial E_i}{\partial A_i} = \sigma \,, \quad p_i \equiv \frac{\partial E_i}{\partial V_i} = p \,, \tag{3.27}$$

as the conditions for stationarity.[15] Thus, the coexistence condition for the members of a given multiplet is that they share same values of σ and p. When all members of the multiplet satisfy $\partial^2 E_i/\partial A_i^2 > 0$ and $\partial^2 E_i/\partial V_i^2 > 0$, then the stationary point is guaranteed to be a local minimum.[16] If either second derivative is negative for all members of the multiplet, then the stationary point cannot be a local minimum.[17] In general, the signs are mixed and nothing can be said without specific calculation.

[15] We have assumed here that A_i and V_i are independent variables. This is true for all Euler shapes except spheres, which require special treatment, since the spherical radius R_i determines both A_i and V_i . We shall comment further in Section 3.3. The conclusion is that this result holds equally for spheres.

[16] An example here is the situation for spheres, which we shall study in detail in Section 3.3. Of course, in the spherical case A and V are not independent and the relevant condition is $\partial^2 E/\partial R^2 > 0$.

[17] An example of this is elliptical shapes which are not too far from spheres, for which $\partial^2 E/\partial A^2 < 0$, as we shall show in Section 3.3.

The stationary point or points given by (3.27) define a set of multiplet energy sheets $E_m(A,V)$ which must be combined with the Euler sheets of Subsection 3.2.1 in selecting the true lowest-energy state (3.23) at each given point (A,V).

3.3 Spheres, sphere multiplets and perturbations about spheres

3.3.1 Spheres, sphere multiplets and vesiculations

In the (A,V) plane, spheres lie along the line $V = A^{3/2}/3\sqrt{4\pi}$, which bounds the physical region from above. Each point on this line corresponds to a particular radius R. The spherical symmetry of the problem guarantees that such shapes will be solutions of the Euler problem and, indeed, we see that $C_m(r) = C_p(r) = 1/R$ satisfies, e.g., Eq. (3.5), provided

$$p = \frac{1+2\sigma}{R} - \frac{2}{R^2}\,, \tag{3.28}$$

which is just the usual soap-bubble formula corrected for the contribution coming from the curvature energy. Note that Eq. (3.28) is linear in both σ and p, so that each sphere is represented by a straight line in the (σ,p) plane. Inverting Eq. (3.28) to find the allowed radii at fixed (σ,p) gives

$$R_{\pm}(\sigma,p) = \frac{(1+2\sigma) \pm [(1+2\sigma)^2 - 8p]^{1/2}}{2p}\,, \tag{3.29}$$

so that spheres are forbidden above the parabola

$$p = (1+2\sigma)^2/8\,. \tag{3.30}$$

Along this boundary (Fig. 3.3), the two allowed spheres have equal radii. In the region below this curve, there are two free-energy sheets $\Phi_{S,\pm}(\sigma,p)$ (one above the other) defined for spheres $R_{\pm}(\sigma,p)$, respectively, as

$$\Phi_{S,\pm}(\sigma,p) = 2\pi\,[2 - R_{\pm}(\sigma,p)]^2 + \sigma\, 4\pi R_{\pm}^2(\sigma,p) - p\, 4\pi R_{\pm}^3(\sigma,p)/3\,. \tag{3.31}$$

The two sheets merge smoothly along the boundary (3.30).

All this structure maps in the (A, V) plane onto the line $V = A^{3/2}/3\sqrt{4\pi}$ with corresponding energy,

$$E(R) = E_0(R) \equiv 2\pi(2 - R)^2 \,. \tag{3.32}$$

The mechanism for this reduction is a simple property of the Legendre transform [55]: The variational free energy for a sphere of radius R has the form,

$$\Phi_0(\sigma, p; R) \equiv E_0(R) + \sigma\, 4\pi R^2 - p\, 4\pi R^3/3 \,, \tag{3.33}$$

which defines a plane $P(\sigma, p; R)$ over the (σ,p) space. Note that the entire plane $P(\sigma, p; R)$ maps under Legendre transformation onto a single point $(4\pi R^2, 4\pi R^3/3)$ in the (A, V) plane. Now, the intersection of the spherical sheet $\Phi_S(\sigma, p)$ with this plane lies along the locus (3.28). Furthermore, because of the relation (3.28), the sheet and the plane meet tangentially along this line. Thus, the entire line of this intersection maps, as part of the plane, into the single point $(4\pi R^2, 4\pi R^3/3)$ in the (A, V) plane.

Many-sphere multiplets are a special case of Eqs. (3.25) and (3.26) but require special treatment, since A_i and V_i are related. For $M = 2$, the partition of area and volume is completely determined by the constraints (3.26), and there is a unique solution (up to an overall interchange $1 \leftrightarrow 2$) provided that $\Delta \leq 2^{1/3} - 1$, where Δ is the fractional excess (total) area defined by $A = 4\pi R_V^2(1 + \Delta)$. Since the (σ, p) lines representing two spheres always cross, the intersection point defines the coexistence condition for the two spheres of radius R_1 and R_2,

$$\sigma = (1/R_1 + 1/R_2) - 1/2 \,, \;\; p = 2/(R_1 R_2) \,. \tag{3.34}$$

When $M \geq 3$, the stationarity condition $\partial(E + \lambda_A A - \lambda_V V)/\partial R_i = 0$ results in a quadratic equation for the multiplet radii, so the conclusion is that the spheres belonging to a multiplet can have at most two different radii. Such states are always local minima with respect to changes of the radii, since the second derivative of Eq. (3.32) is positive. When $V \gg 1$ and Δ is not too large, then the lowest-energy sphere-multiplet always uses one large sphere and $n = 1, 2, 3...$ small ones. Representative sphere-multiplet energy diagrams are shown in Fig. 3.4 for fixed volumes, $V = 272$ and $V = 2723$. Near $\Delta = 0$ there exist other non-multiplet configurations of lower energy; however, these multiplets always provide one route

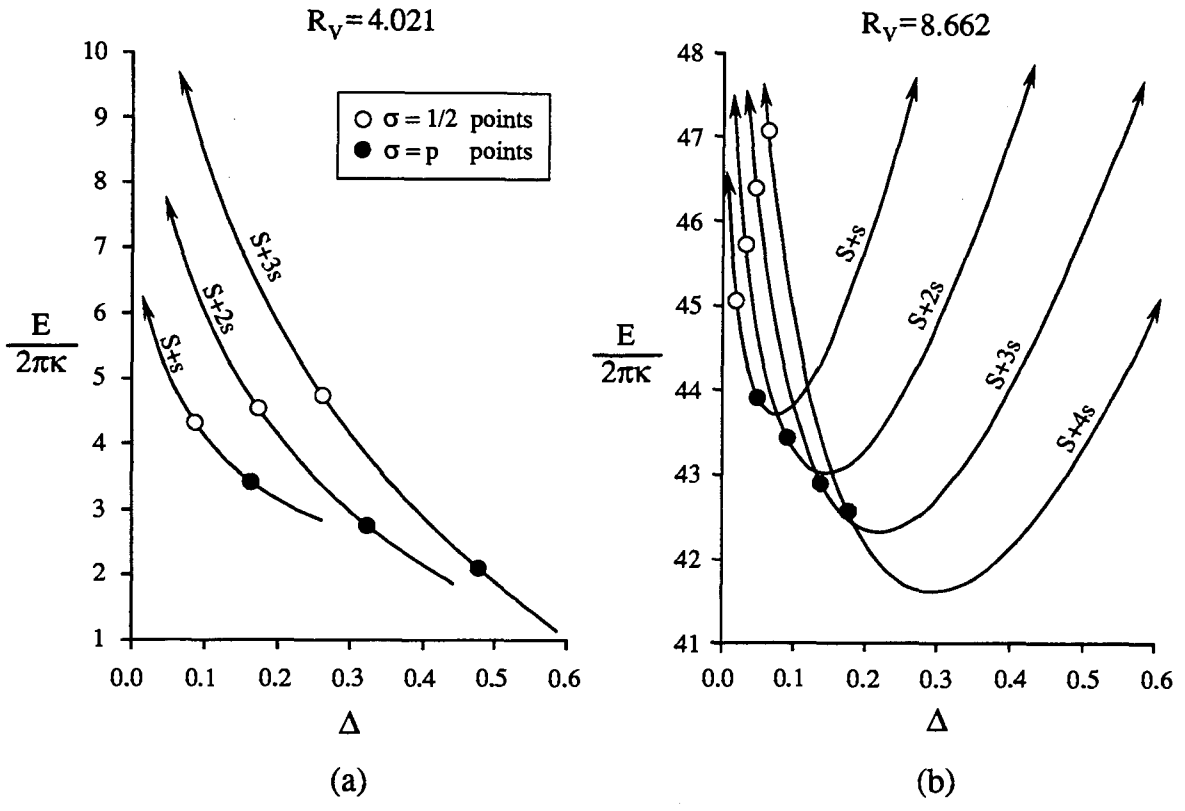


Figure 3.4: Curves of energy E versus fractional excess area (Δ) for sphere multiplets (S+ns) of given volume. The largest possible Δ of (S+ns) occurs when all spheres have the same radius, i.e., at $\Delta = (n+1)^{1/3} - 1$, independent of volume. Two typical cases are shown. (a) For small volume ($R_V = 4.02$), the curves do not intersect, and continuity of the energy requires additional phases. (b) For larger volume ($R_V = 8.66$), successive curves intersect (additional phases are required here but for more subtle reasons). At the special points $\sigma = p$ and $\sigma = 1/2$, spheres in the multiplet satisfy the kissing condition. Note that the $\sigma = 1/2$ point is only relevant (i.e., lowest energy) for the first multiplet. At larger volume, the $\sigma = p$ points become irrelevant for high multiplets.

whereby the original sphere may lower its energy as its area is increased. We shall discuss in the following two sections the extent to which, starting from large initial volumes, this route is actually followed. Note that, when it is, the transitions are first order. When V is sufficiently large, there are no gaps early in this process; for $V \leq 420$ a gap first occurs between $n = 1$ and $n = 2$, so other configurations are necessarily required.[18]

Properties (b) and (c) summarized in the preceding section (for more details, see Appendix C) give some further insights. There are generically two types of kissing boundaries at which the kissing condition (3.24) can be satisfied, as illustrated in Fig. 3.5:

(i) When the two spheres involved are of equal radius, Eq. (3.24) is only fulfilled when $R_A = R_B = R_0 = 2$ (Helfrich spheres), i.e., along the line $\sigma = p$. Sphere-multiplet configurations ($n \geq 2$) near this line are always close in energy to shapes consisting essentially of Helfrich spheres connected through narrow necks (dumbbells or multiple dumbbells). The local stability near this kissing boundary depends on the sign of the coefficient C, or, explicitly, the sign of the bracket $(1 - 2/R_A)$. It follows from Eq. (3.29) that, for $\sigma > 1/2$, the Helfrich spheres are given by the solution $R_+(\sigma, p)$ (the larger of the two equilibrium radii) at $\sigma = p$. It is easy to see that the bracket $(1 - 2/R_+(\sigma, p))$ is positive to the right of the line $\sigma = p$ and negative to the left. The situation reverses when $\sigma < 1/2$. The Helfrich spheres now correspond to the solution $R_-(\sigma, p)$ at the line $\sigma = p$. In this case, the bracket is given by $(1 - 2/R_-(\sigma, p))$, which turns out to be negative to the right and positive to the left of the $\sigma = p$ line. For $0 < \sigma < 1/2$,[19] thus, the stable (i.e., low-energy) configuration will be dumbbells to the right and sphere-multiplets to the left. Similar statements can be made for the case in which $\sigma > 1/2$.

(ii) When the two spheres have unequal radii, the kissing condition is satisfied when $\sigma = 1/2$ (and[20] $p < 1/2$). Near this line, we expect competition from asymmetric dumbbell shapes (or multiple dumbbells involving alternate sizes). The asymmetric necks are stable for $\sigma > 1/2$ and the multiplets, for $\sigma < 1/2$. These special points are indicated in Fig. 3.4

[18] Actually, for reasons discussed in Section 3.5, these other configurations come in slightly before the gap opens.

[19] For $\sigma < 0$, the stability is expected to be different. See discussion in Appendix C.

[20] For $p < 0$, one of the spheres, $R_-(\sigma, p)$, acquires a negative radius. The vesiculation is then inverted ("invagination").

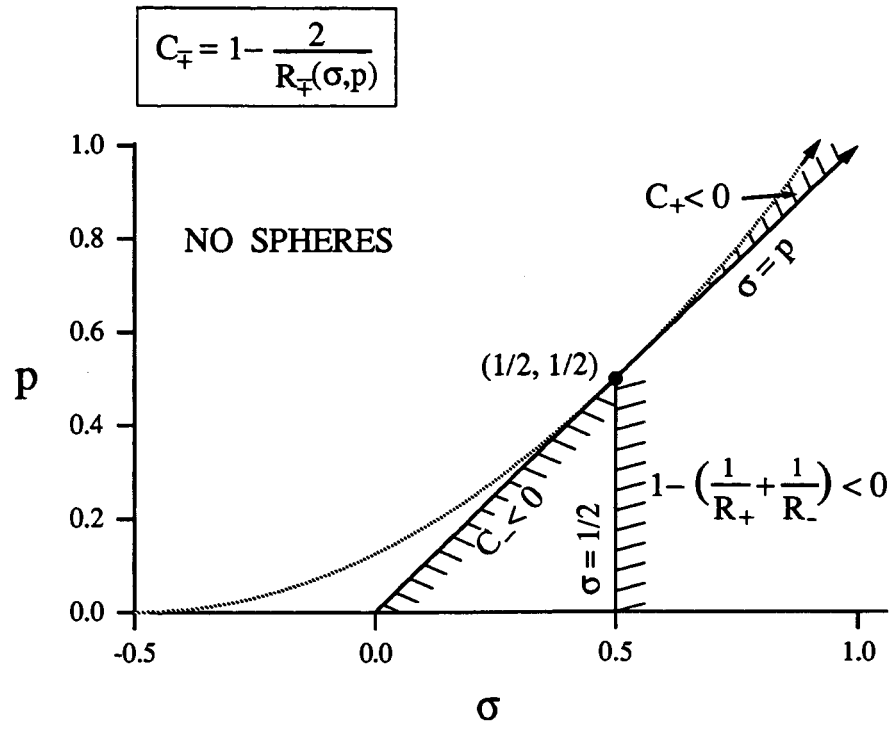


Figure 3.5: The kissing boundaries of sphere multiplets and the local stability of small necks. Notation is similar to that used in Fig. 3.3. The two kissing boundaries, ($\sigma = p$) and ($\sigma = 1/2$), are shown by solid lines. Along ($\sigma = p$), the kissing spheres are all Helfrich spheres; along ($\sigma = 1/2$), the coexisting spheres have distinct radii (Eq. (3.29)). The hatched regions represent the regions where small necks are stable.

and will play an important role in later discussion. Note that the $\sigma = 1/2$ point is only relevant (i.e., low-energy) for the smallest multiplet configuration, (S+s), which consists of two spheres of unequal radii. On the other hand, for V sufficiently large, the $\sigma = p$ points remain lowest-energy configurations for most higher multiplets (S+ns), as we shall see in Section 3.5.

3.3.2 Perturbations about spheres

We now turn to the discussion of nearly spherical shapes. The discussion illustrates the continuous development of quasi-spherical shapes from spheres ("bifurcation"), as parameters are varied, and also leads to certain analytical results useful to the determination of

the final phase diagram.

Nearly spherical shapes can be treated perturbatively [51, 52, 53, 54], provided that the fractional excess area Δ is not too large. For axisymmetric shapes S, we may expand in spherical coordinates in terms of the $m = 0$ spherical harmonics,

$$R(\theta) = R\left\{1 + \sum_{l=1}^{\infty} a_l Y_{l,0}(\theta)\right\} . \tag{3.35}$$

The functionals $E[S]$, $V[S]$, $A[S]$ and $\Phi[S]$ all become functions of the distance R and the coefficients $\{a_l\}$. Because of Euclidean invariance,[21] there is no loss of generality in setting $a_1 = 0$. Thus, for example,

$$\begin{aligned} \Phi(\sigma, p; R, \{a_l\}) &= \Phi_0(\sigma, p; R) + \sum_{l=2}^{\infty} C_l^{(2)}(\sigma, p; R)\, a_l^2 \\ &\quad + \sum_{l_1, l_2, l_3}^{\infty} C_{l_1, l_2, l_3}^{(3)}(\sigma, p; R) a_{l_1} a_{l_2} a_{l_3} + \cdots , \end{aligned} \tag{3.36}$$

with

$$\begin{aligned} C_l^{(2)}(\sigma, p; R) &= \left\{\frac{1}{2}(l-1)l(l+1)(l+2) - Rl(l+1)\right. \\ &\quad \left. + R^2(\frac{1}{2} + \sigma)\left[1 + \frac{l(l+1)}{2}\right] - pR^3\right\} . \end{aligned} \tag{3.37}$$

Terms of order a^3 and a^4 are given in Appendix D, which provides a complete treatment of the local bifurcation structure using the perturbation expansion about spheres. For given σ and p, the stationarity of $\Phi[S]$ with respect to R and a_l leads to a set of nonlinear equations equivalent to, e.g., Eq. (3.5). The condition $\partial\Phi/\partial R = 0$ leads to $R = R_0(\sigma, p) + O(a_l)$ with $R_0(\sigma, p)$ given by Eq. (3.28). The condition $\partial\Phi/\partial a_l = 2C_l^{(2)}(\sigma, p; R)a_l + \cdots = 0$ shows that, for each integer $l > 1$, there is a sheet of nearly spherical Euler shapes which bifurcates continuously away from spheres along the line $C_l^{(2)}(\sigma, p; R_0) = 0$. The locus of this bifurcation is given parametrically [53] by

$$\sigma = \sigma_l = -\frac{1}{2} + 2/R_0 - l(l+1)/R_0^2 ,$$

[21] E, Φ, A, and V are invariant under Euclidean transformations of S. What is important here is that translations of S along the symmetry axis induce nonlinear transformations on the coefficients $\{a_l\}$ but leave the values of these functionals unchanged. By means of such a translation it is always possible to make $a_1 = 0$.

$$p \;=\; p_l = -2[l(l+1) - R_0]/R_0^3 \,, \tag{3.38}$$

as illustrated in Fig. 3.3. The stationarity conditions, $\partial\Phi/\partial R = 0$ and $\partial\Phi/\partial a_l = 0$, are equivalent to the Euler-Lagrange equations given in Section 3.2.1. The perturbative solution to these variational equations corresponds to a quasi-spherical Euler shape. Its evolution with the change of σ and p leads to a free-energy sheet which separates smoothly from the lower of the two sphere sheets along each of these boundaries, as illustrated schematically in Fig. 3.6.

The mechanical bending energies of these perturbative quasi-spherical shapes, as functions of A and V, are of essential importance to constructing the phase diagram. In order to develop these energies, we now make use of the fact that these perturbative solutions, when treated in the (A,V) ensemble, can be developed in powers of $\Delta^{1/2}$. It turns out [52] that the $l = 2$ prolate ellipse sheet provides the lowest-energy shape among all the deformed spheres characterized by different $l's$, when the fractional excess area (Δ) is small. Consider a prolate ellipse which bifurcates from a sphere of radius R_V via the $l = 2$ mode, and preserving the volume of the sphere but having an excess area of Δ. Expansion near the bifurcation takes the form,[22]

$$\begin{aligned} R &= R_V[1 - \Delta/2 + O(\Delta^{3/2})] \\ a_{2n+1} &= 0, \;\; a_2 = \sqrt{2\pi}\Delta^{1/2} + O(\Delta), \;\; a_{2n} = O(\Delta^{n/2}) \quad . \end{aligned} \tag{3.39}$$

The energy of this shape is given by

$$[E(\Delta, R_V) - E_0(R_V)]/2\pi = (R_V^2 - 4R_V + 12)\Delta + O(\Delta^{3/2}) \,, \tag{3.40}$$

which shows how the energy increases as the area increases at constant volume. Terms through $O(\Delta^2)$ are given in Appendix D and will be needed in Section 3.5, as we consider budding and vesiculation in the large-volume limit.

[22]The absence of odd l terms is the general consequence of the up/down symmetry of the shape in the perturbative regime.

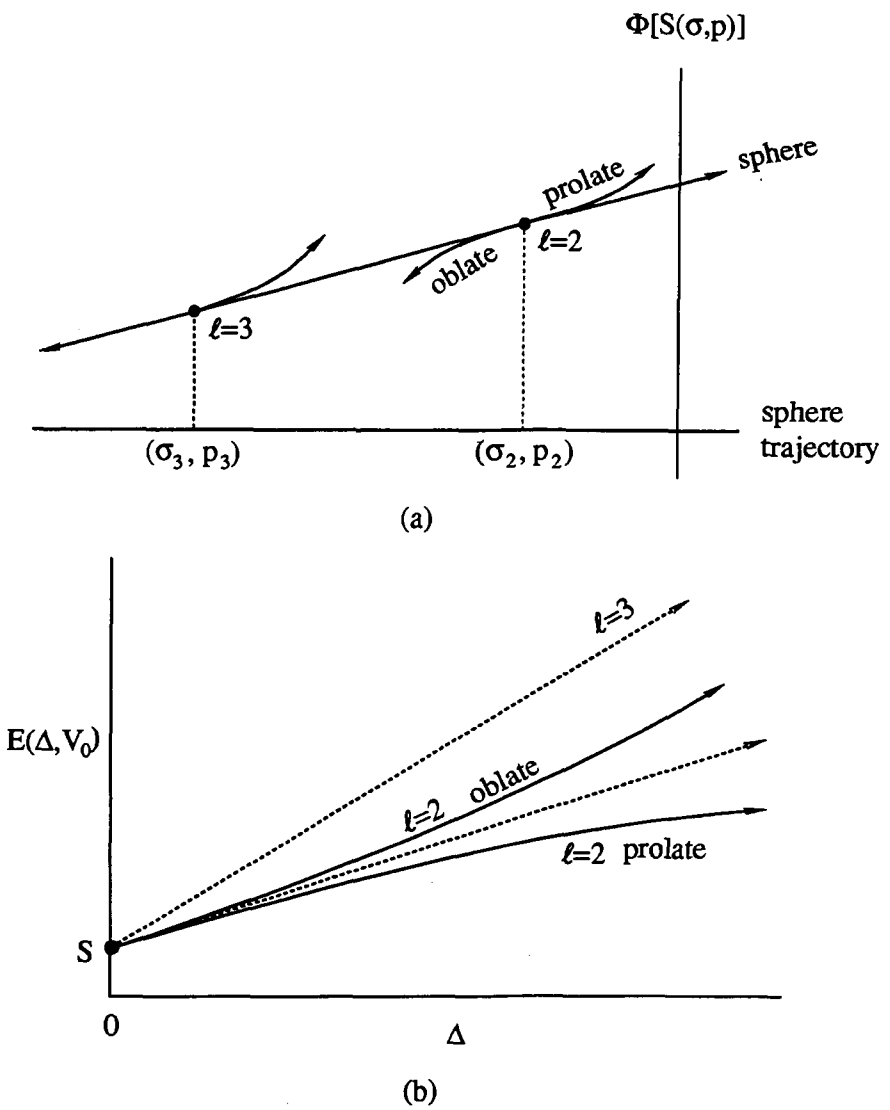


Figure 3.6: Structure of bifurcations from spheres. (a) A schematic view of a section of the free-energy sheets Φ_E and Φ_S, cut along a straight line in the (σ, p) plane corresponding to a sphere of fixed radius R. Bifurcation of the $l = 2$ and $l = 3$ sheets away from the sphere sheet is characterized by the separations between the sheets, $\|\sigma - \sigma_2\|^3$ and $\|\sigma - \sigma_3\|^2$, respectively. Note, also, that the $l = 3$ bifurcation is one sided (see Appendix D). (b) The same structure in the (A, V) plane, showing how the two bifurcation points coincide in this representation. The lowest-energy sheet near $\Delta = 0$ is always the prolate branch of the $l = 2$ shape. The dashed lines show the leading corrections to the sphere energy, linear in Δ. The next correction go as $\Delta^{3/2}$ for $l = 2$.

3.4 Shape evolution at $V = 272$: numerical results

To set the stage for a more general discussion of the shape evolution and phase diagram at a large (fixed) volume as area $A = 4\pi R_V^2(1+\Delta)$ increases away from the sphere, we present in this section results for a volume $V = 272$, i.e., $R_V = 4.02$ (in units such that $C_0 = 1$). The numerical algorithms used for the shape calculations on which these results are based are explained in Appendix B. In short, the Euler-Lagrange equations (3.8)–(3.12) were integrated by Runge-Kutta methods. The numerical integration becomes rather unstable for large volumes and the choice in our example represents a volume which is appreciably larger than unity but still in a domain where numerical convergence is excellent and details of neck shape can be studied reliably.

Figures 3.7 and 3.8 illustrate the shape evolution, including both the energy as a function of Δ and sketches of the corresponding minimizing shapes. Although the energy is continuous as a function of Δ, the curve consists of many segments and is certainly not convex. Each segment corresponds to a region or "phase" in the (A, V) space, where the shape evolves smoothly. The segments are separated by phase transitions. Both first-order transitions (slope discontinuity) and second-order phase transitions (no slope discontinuity) are represented. The initial ($\Delta = 0$) spherical shape (S) immediately passes through the $l = 2$ threshold (Section 3.3 and Appendix D) into a prolate elliptical shape (E), which persists until fractional excess area $\Delta = 0.07$. At this point (see inset on Fig. 3.7, point 1) there is a first-order budding transition (in which up/down symmetry is broken) to a necked phase (N) with small neck radius ($a = 0.03$). As Δ is increased further, the neck radius shrinks smoothly to zero at $\Delta = 0.086$, where there is a second-order vesiculation transition (point 2) to a multiplet (S+s) consisting of one large sphere ($R_1 = 3.97$) plus one sphere with a radius ($R_2 = 1.34$) somewhat smaller than the Helfrich sphere. On further increase of the area, the smaller sphere grows at the expense of the larger one, until at $\Delta = 0.260$ the two spheres have equal volume. At this point (3), one of the two spheres deforms continuously into an elliptical shape, having a volume which grows with Δ, and the remaining sphere starts to shrink (E+s). The transition (S+s) $\rightarrow$(E+s) involves a discontinuous jump in the

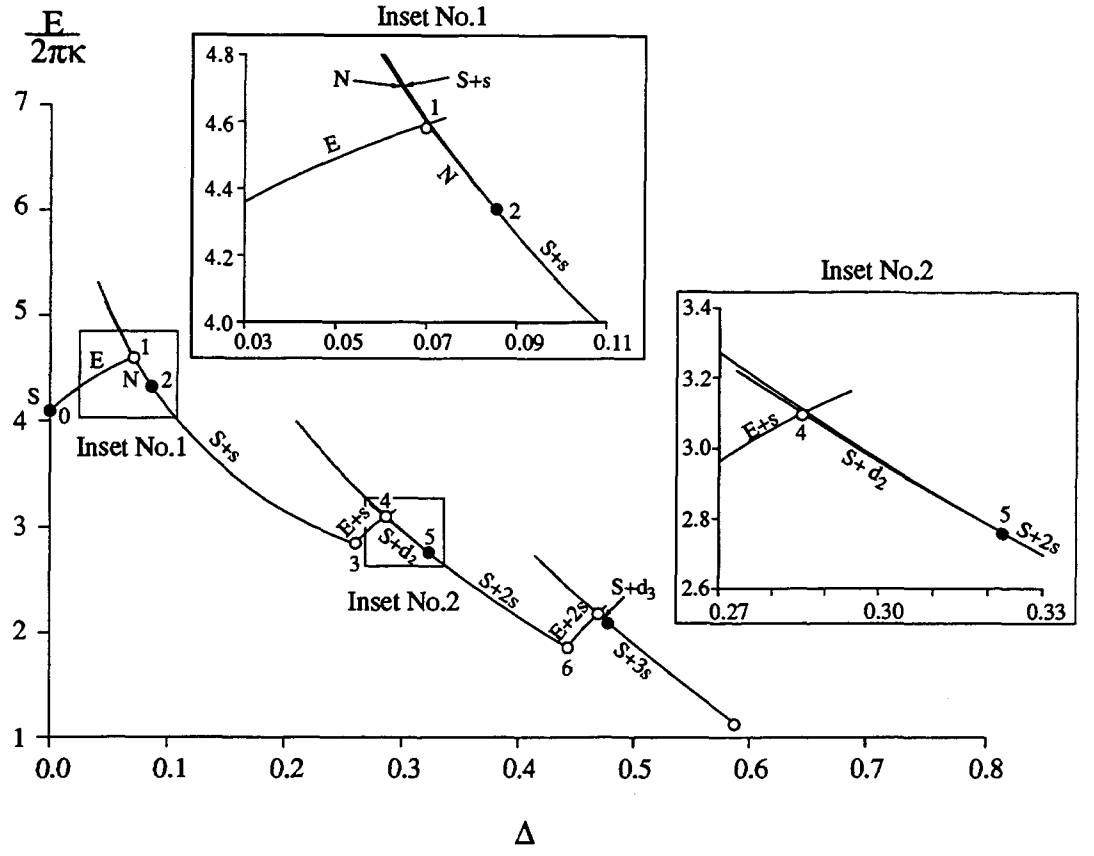


Figure 3.7: Energy versus fractional excess area Δ for $R_V = 4.02$ ($V = 272$). The successive phases (S)→(E)→(N)→(S+s)→ $\cdots$ are explained in Fig. 3.8. Insets show extra detail in special regions. The sequence from (S+2s) through (S+3s) parallels that of (S+s) to (s+2s), shown in the inset. Phase boundary points are labeled S(0),1,2,.. and appear again in the (σ, p) diagram, Fig. 3.9. First-order transitions (slope discontinuity) are shown as open circles; second-order points (no slope discontinuity) are shown as solid circles. The data are partially analytic and partially from numerical solution of the Euler equation.

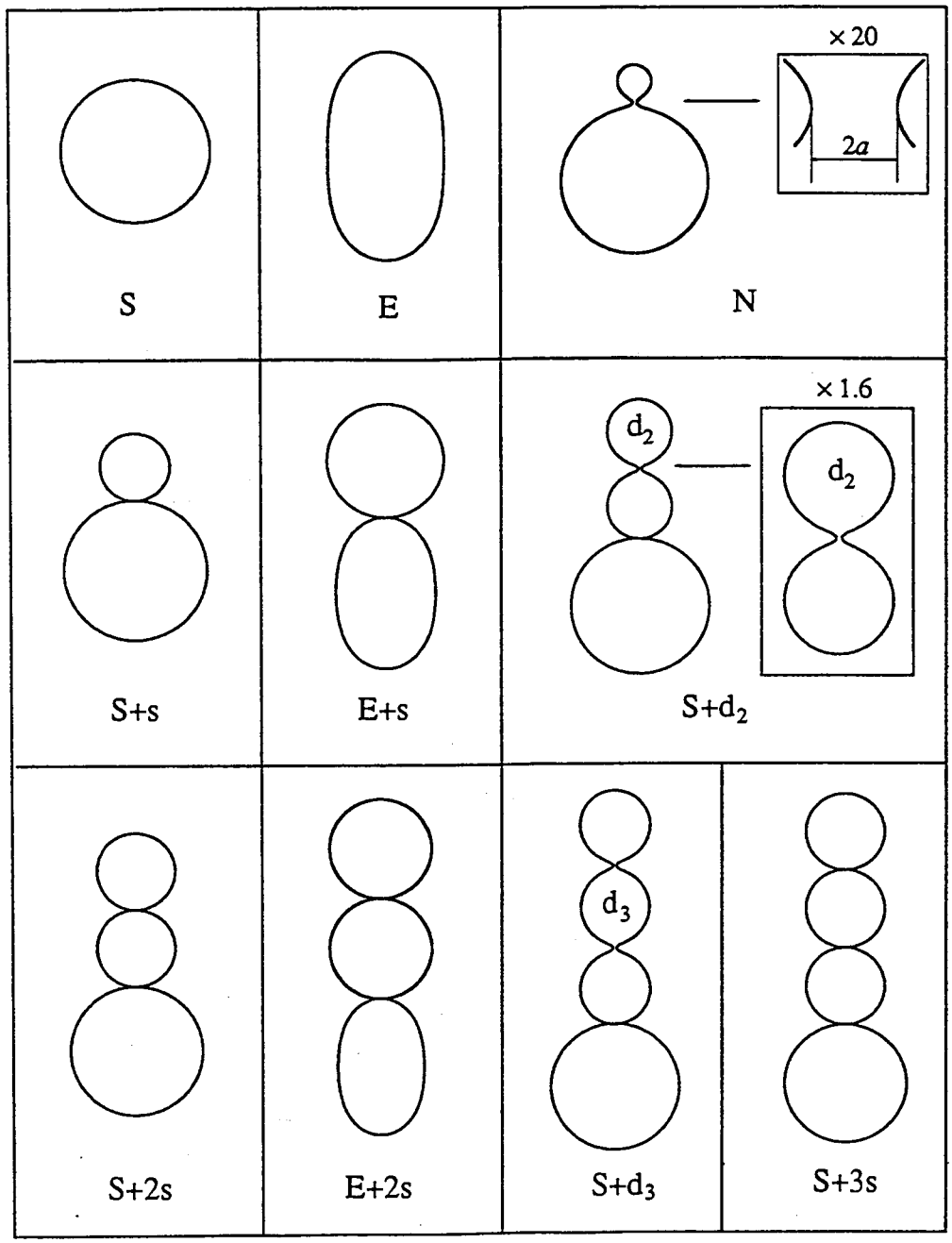


Figure 3.8: Shape evolution of a vesicle as Δ increases, corresponding sequentially to the different regions (phases) of Fig. 3.7 ($R_V = 4.02$). S represents sphere; E,"ellipse"; N, neck; S+s, two-sphere multiplet; E+s, "ellipse" plus sphere multiplet; S+d_2, sphere plus two-part dumbbell; etc.

(σ, p) plane and is first order, despite the fact that the actual shape changes are continuous. The (E+s) energy curve crosses the (S+2s) energy curve at $\Delta = 0.2866$; however, in this region the (S+2s) multiplet is unstable to another multiplet, (S+d_2), in which a large sphere coexists with a small symmetric dumbbell. Thus, the actual transition is (E+s) $\rightarrow$(S+d_2) at $\Delta = 0.2857$ (first order, point 4), followed by (S+d_2) $\rightarrow$ (S+2s) at $\Delta = 0.327$ (second order, point 5) on the $\sigma = p$ line, where the neck of the dumbbell shrinks to zero. This scenario appears to continue, with (S+2s)$\rightarrow$ (E+2s)$\rightarrow$ (S+d_3)$\rightarrow$ (S+3s)$\rightarrow$ ..., where d_3 represents a three-component symmetric dumbbell. The only uncertainty is whether at some stage the $\sigma = p$ point moves far enough to the left so that the transition (E+ns)$\rightarrow$ (S+$(n+1)$s) takes place directly. Of course, other shapes presumably appear at sufficiently large Δ; we have only followed the process through $\Delta = 0.5$.

The trajectory of the shape evolution in the (σ, p) plane is shown in Fig. 3.9. Second-order transitions (solid dots) are continuous in this representation; however, first-order transitions involve a discontinuous jump, so each open circle appears twice and there would be corresponding tie lines on the full (σ, p) phase diagram. The bifurcation lines $\sigma = 1/2$ and $\sigma = p$ are shown for convenience, along with both the $l = 2$ threshold (see Section 3.3) and the boundary, Eq. (3.30), of the regions allowed for sphere solutions. Note, in particular, the mechanism for the (S+s)$\rightarrow$(E+s) transition: The (S+s) phase ends at the boundary of the allowed-sphere region, where the radii of the two spheres are equal at $R_1 = R_2 = R_A/\sqrt{2}$. The phase point then jumps along the (σ, p) line representing a sphere of this radius to its intersection with the $l = 2$ threshold. This explains how the transition can be first order but still have a smooth shape variation.

We close this Section with a remark on rigor. The Euler equations are nonlinear and have many solutions. It is not hard to follow a particular solution numerically as σ and p are varied; however, there is no general method for proving whether or not the branch one is following has the lowest energy. The existence of multiplet regions makes this problem even more acute, since, in principle, one must test all possible multiplet combinations to search for the lowest energy. We cannot claim to have done this in any rigorous manner. The phase diagrams shown in this and the next sections are the result of a search only over

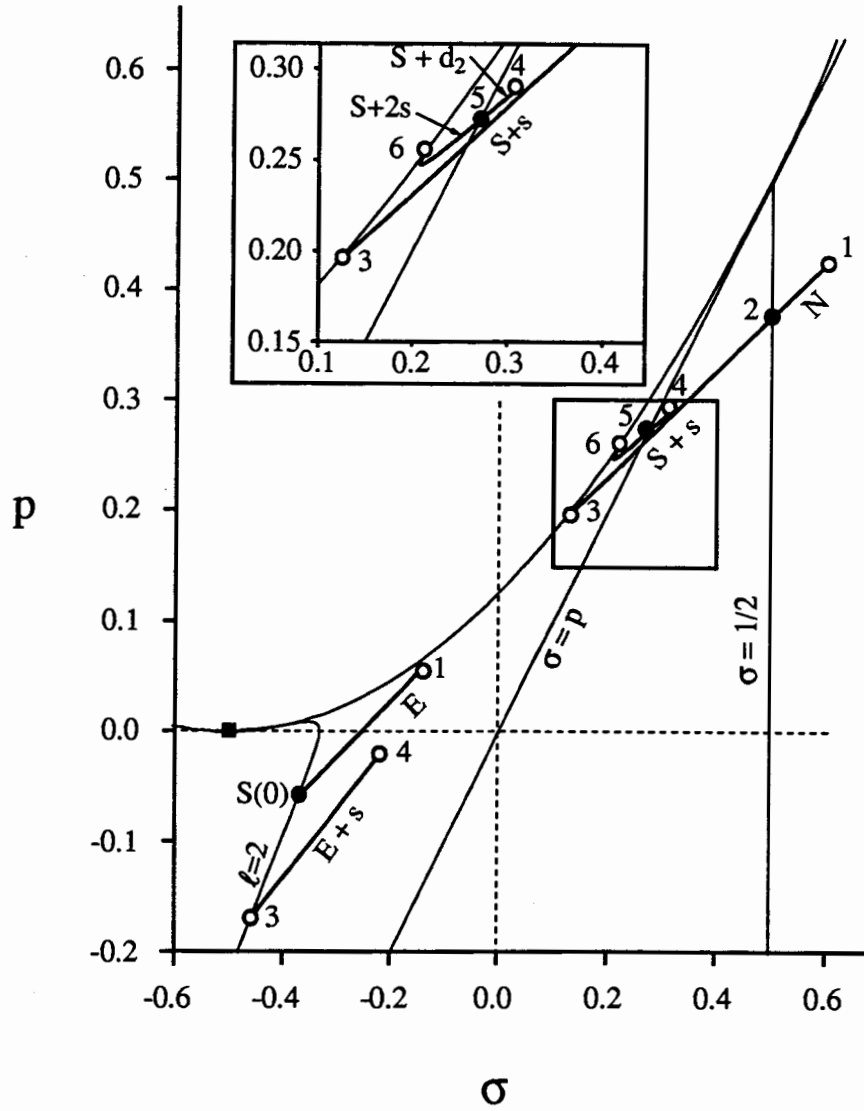


Figure 3.9: (σ, p) trajectory of the shape evolution for $R_V = 4.02$. The shape evolution in Fig. 3.8 is traced out in the (σ, p) plane. Notation corresponds to that of Fig. 3.7 and Fig. 3.8. Note that a first-order transition corresponds to a discontinuous jump in the (σ, p) plane, so that the corresponding point appears twice (e.g., point 1). Second-order transitions occur at bifurcation lines (here, $l = 2$, $\sigma = p$, and $\sigma = 1/2$, shown as thin lines). Inset shows detail.

those configurations, Euler shapes and multiplets, which we have found and considered likely candidates.

3.5 Analytical results for the phase diagram at large volume

To find the full phase diagram, one may repeat the calculations of Section 3.4 for many different volumes and simply tabulate the results. A rather complete study of this sort has recently been carried out by Seifert *et al.* [27] for a range of $R_A, R_V \sim 1$ (in $C_0 = 1$ units). On the other hand, as we shall outline in this Section, at large volumes and fractional excess areas which are not too large, the phase boundaries may be mapped out quasi-analytically. In this region the numerical integrations become unstable, and, in this sense, our work complements that of Seifert *et al.*, extending it to large volumes, where budding and vesiculation dominate the phase diagram.

Conceptually, the scenario is as anticipated in Section 3.3: A sphere of volume NV_H with $N \gg 1$ ($V_H = 32\pi/3$ is the volume of the Helfrich sphere) vesiculates off a sequence of small spheres, as its area is increased. As long as $A \ll NA_H$ (where $A_H = 16\pi$ is the area of the Helfrich sphere), the minimum-energy configuration is always a multiplet which consists in effect of n ($\ll N$) Helfrich spheres (which cost no energy) and one larger sphere which carries the remaining volume. As area increases, the energy decreases from $E = 8\pi(N^{1/3} - 1)$ at $A = 4\pi R_V^2$ to near zero at $A = N\,A_H$ ($E = 0$ when N is an integer).

These successive vesiculations are essentially periodic in A with period A_H, as each increment of area is shed in the form of an approximate Helfrich sphere. Between successive vesiculations, the additional area can be absorbed either by increasing the radius of the small spheres at the expense of the large one or by deforming the large sphere. The former route is always preferable when it is possible; but, as we shall see below, it is not always possible at low volume. The vesiculation $(n-1) \rightarrow n$ normally occurs when the radius of the n small spheres is slightly less than R_0, i.e., at a point just to the right of the line $\sigma = p$, where, as discussed in Section 3.3 and Appendix C, $C < 0$ and the n-fold multiple dumbbell d_n exists as an Euler shape (see Fig. 3.10). The multiplet phase (S+d_n) always has lower

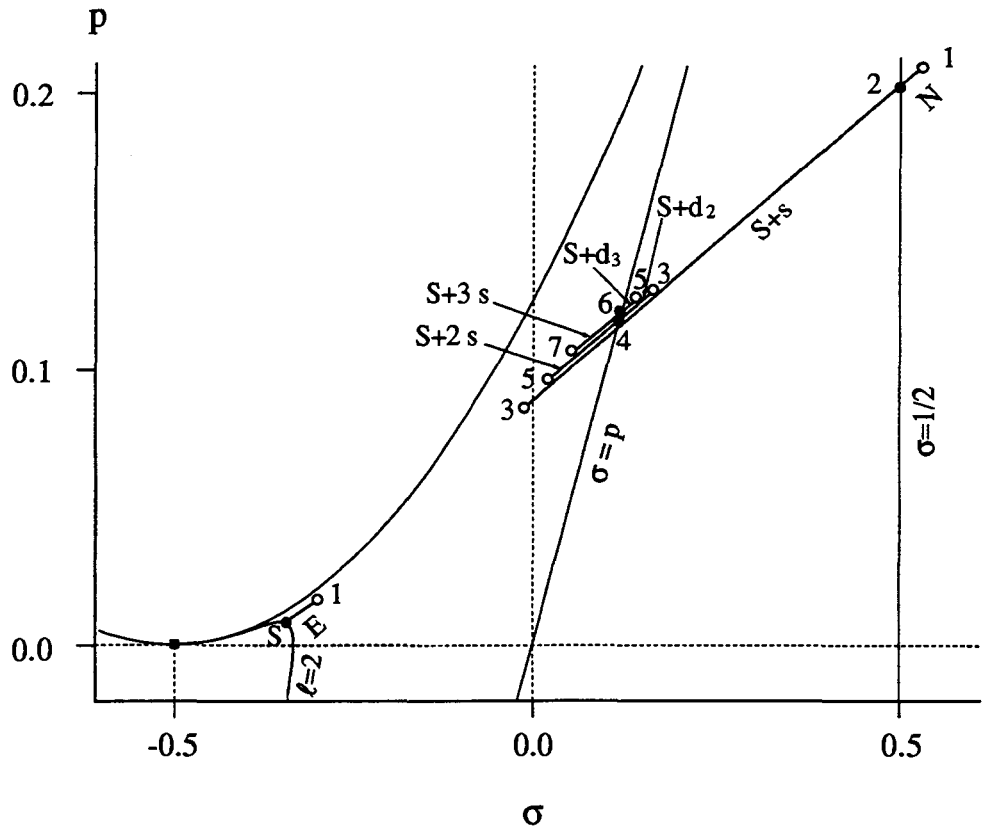


Figure 3.10: (σ, p) trajectory of the shape evolution for $R_V = 8.66$ ($V = 2723$). Same as Fig. 3.9 but for the larger volume $R_V = 8.66$ ($V = 2723$). Note here that, in contrast to Fig. 3.9, the multiplet phase (S+s) becomes unstable to (S+d_2) before reaching the parabolic boundary. (S+d_2) transforms to (S+2s) at a second-order transition, when it reaches the bifurcation line $\sigma = p$. For nonspherical shapes the data are approximate, based on perturbative results. Separation between multiplet trajectories are exaggerated by about a factor of 2 in order to make the structure visible at this scale.

energy than the (S+ns) phase. The dumbbell phase (S+d_n) therefore precedes the phase (S+ns) in the vesiculation sequence. The scenario sketched above holds until n becomes comparable to N. For $n \gtrsim N$ (large area A), the bending energy E must begin to increase with A. At this point, there is much excess area available and it becomes energetically preferable to use nonspherical shapes in the multiplets, so other configurations may occur. This large-excess-area, large-volume region of the phase diagram remains to be explored.

Figure 3.11 shows the large-volume phase diagram, which summarizes our calculational results and includes data through the fourth vesiculation. The lowest volume shown corresponds to $V = 2V_H$ ($R_V = 2^{4/3}$), which is the minimum volume at which the first vesiculation can take place via the mechanism we have described, since the $\sigma = \frac{1}{2}$ bifurcation line (Fig. 3.4) ends at $p = \frac{1}{2}$, where the corresponding radius is that of the Helfrich sphere. Volumes comparable to and smaller than this have been explored in Ref. [27].[23] More generally, when $R_V/R_H \sim n^{1/3}$, new phases and new phase sequences–not shown in Fig. 3.11–can and do occur. Generally, the (S + (n − 1)s) phase does not occur below the point labelled H_n with coordinates [$\Delta = n^{1/3} - 1$, $R_V = 2n^{1/3}$], where n Helfrich spheres coexist.

The sequence of phases, as Δ increases at fixed V, is similar to what we observed in Section 3.4, but with some important modifications:

(a) One difference is the absence above appropriate critical volumes of the (E+ns) phases, which in Section 3.4 intervened between (S+ns) and (S + d_n) in the sequence (S+ns) → ... → (S + (n + 1)s). The reason for this–already anticipated in Fig. 3.4–is that the energy curves $E_{\mathrm{S}+n\mathrm{s}}(\Delta)$ overlap at sufficiently high volume, and the small-sphere (s) members of the multiplets remain on the scale of the Helfrich sphere. In Fig. 3.11 we have denoted by E_{n+1} the largest-volume point of the phase (E+ns).

(b) Another difference is the disappearance of the dumbbell phases (S + d_n) at intermediate volumes for $n \geq 3$. The mechanism for this disappearance is the movement of the $\sigma = p$ point into a physically irrelevant (i.e., high-energy) region over this interval. The

[23] According to Ref. [27], just below $V = 2V_H$ the sequence of phases is reentrant: (E) goes to (N) (called "pear" in Ref. [27]) via a first-order transition, but then the neck grows again and, at a higher area, (N) reconverts to (E) by a bifurcation (second-order).

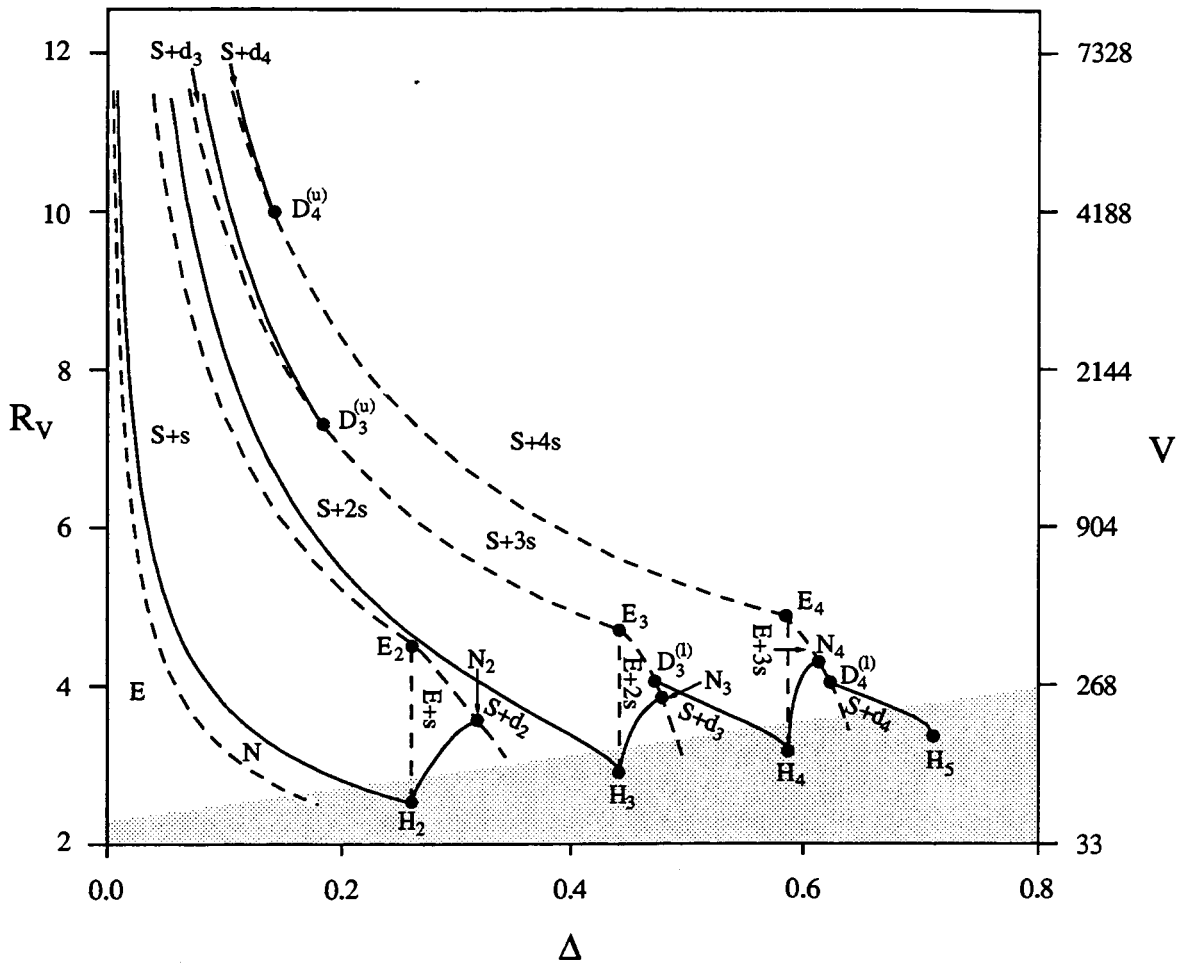


Figure 3.11: Phase diagram for the spontaneous-curvature model in the region of large volume [$V = 4\pi R_V^3/3$] and small excess area [$A = 4\pi R_V^2(1+\Delta)$]. The labeling of the phases is given in Fig. 3.8. First-order transitions are shown by dashed lines; second-order transitions, by solid lines. The location of special points is provided in Table I. The gray region corresponds to $\Delta \gtrsim (R_V/2) - 1$ (i.e., areas comparable to and larger than N Helfrich spheres), where the phase diagram is not controlled by the arguments given here. At large excess area, the phase diagram is not yet known, and it is probable that nonaxisymmetric phases play a role.

	Δ	R_V
H_2	$2^{1/3} - 1 = 0.260$	$2 \times 2^{1/3} = 2.520$
H_3	$3^{1/3} - 1 = 0.442$	$2 \times 3^{1/3} = 2.884$
H_4	$4^{1/3} - 1 = 0.587$	$2 \times 4^{1/3} = 3.175$
E_2	$2^{1/3} - 1 = 0.260$	4.65 ± 0.02
E_3	$3^{1/3} - 1 = 0.442$	4.71 ± 0.01
E_4	$4^{1/3} - 1 = 0.587$	4.91 ± 0.01
$D_3^{(u)}$	0.187 ± 0.005	7.22 ± 0.05
$D_4^{(u)}$	0.140 ± 0.005	9.95 ± 0.05
$D_3^{(l)}$	0.472 ± 0.005	4.06 ± 0.02
$D_4^{(l)}$	0.623 ± 0.005	4.04 ± 0.02
N_2	0.318 ± 0.005	3.52 ± 0.05
N_3	0.476 ± 0.005	3.92 ± 0.05
N_4	0.613 ± 0.005	4.30 ± 0.05

Table 3.1: Location of special points in the phase diagram of the SC model (Fig. 3.11). H_n corresponds to a multiplet of n coexisting Helfrich spheres ($R_0 = 2$). E_{n+1} are the points above which the phase (E+ns) does not occur. $D_n^{(u)/(l)}$ are the points above/below which the dumbbell phase (S+d_n) occurs. N_n are the intersection points of the (E+ns)-energy curves with the (S+d_n)-energy curves, at which the kissing condition ($C_E + C_s = 1$) is also satisfied. Where analytical expressions are not shown, the coordinates rely on numerical results and/or perturbative expressions, and corresponding uncertainties are given.

corresponding lower and upper boundaries of the (S + d_n) phase we denote $D_n^{(l)}$ and $D_n^{(u)}$, respectively.

(c) Finally, at volumes near but above $(n+1)V_H$, the (E+ns) phase reaches the (E)-(s) kissing boundary, as area increases, before intersecting with (S + d_{n+1}). At the kissing boundary the condition $C_E + C_s = 1$ is satisfied, so, based on the discussion given in Appendix C, one expects that narrow-necked shapes are close in energy to the vesiculated E+ns and that they may come into play, as do those small-necked shapes near spherical multiplets. However, it does not seem to be true that a small neck is always mechanically stable, as would be the case where the osculating objects are all spheres. The structure

of the transition sequence in this regime may become very complex.[24] We have not investigated this regime thoroughly and shall mark in the phase diagram only the kissing boundaries of the E+ns multiplets, which are represented by the curves between points H_{n+1} and N_{n+1}.

Positions of the special points, H_n, E_n, $D_n^{(l)}$, $D_n^{(u)}$, and N_n, are given in Table I.

We turn now to the location of the phase boundaries. The necking boundary, (E)→(N), is first order and must in principle be calculated numerically. However, the energy curves for the (N) and (S+s) phases lie very close to one another, as is already visible in Fig. 3.7, so an excellent approximation can be obtained simply by equating the perturbative expression (3.40) (including the higher-order corrections (Eq. (D.39)) to the energy of the (S+s) multiplet. At high volume, the fractional excess area at the transition is $\Delta_1 = 1/R_V^2 + O(1/R_V^3)$. The (second-order) vesiculation boundary, (N)→(S+s), is just the $\sigma = \frac{1}{2}$ curve, $R_V = 1/r_1(\Delta)+1/r_2(\Delta)$ with $r_1^3+r_2^3 = 1$ and $r_1^2+r_2^2 = 1+\Delta$. The range of Δ over which the necked phase (N) exists decreases asymptotically in width like $(const.)/R_V^4$ as $R_V \to \infty$.[25] The boundary between (S+ns) and (E+ns) occurs when the radius of the n small spheres has grown to that of the coexisting large sphere, i.e., along the sphere boundary (3.30), which gives $\Delta = (n+1)^{1/3} - 1$, independent of volume. (E+ns) $\leftrightarrow$ (S + d_{n+1}) is first-order [like (E)$\leftrightarrow$(N)] but, again, because the dumbbell neck energy is small, it is well approximated by equating the (E+ns) energy (numerical or perturbative) to the (S + (n + 1)s) energy (analytic). Finding the kissing boundary (E+ns) requires knowing the volume and axial curvature of the "ellipse" shapes, both of which can be well approximated perturbatively, giving the curves shown. Finally, the second-order vesiculation boundary (S + d_n) $\leftrightarrow$ (S+ns) is the curve $\sigma = p$, $R_V = 2/r_2(\Delta)$, with $r_1^3 + r_2^3 = 1$ and $r_1^2 + nr_2^2 = 1 + \Delta$.

[24] Involved in the competition for the lowest energy shape are not only the multiplets E+ns, and nearby small-necked shapes, but also those necked shapes bifurcating from other multiplets such as m_1E+m_2s where $m_1 + m_2 = n + 1$, although these multiplets, themselves, are higher-energy shapes and are excluded from the competition.

[25] Interestingly, the value of the constant seems not to be fully determined by the coefficients in Appendix D; however, we believe that the evidence is strong that it is positive and of order unity.

Chapter 4

Budding and Vesiculation in the ADE Model

4.1 Introductory remarks

The motivation to study the area-difference-elasticity or ADE model is the failure of either of the two earlier models (SC and ΔA) to consistently describe available experiments, along with the fact (emphasized in Chapter 2) that the global elasticity term–omitted by these models–is generically of the same order of magnitude as the local bending energy. Our study of the SC model is described in Chapter 3 and Ref. [29]. A parallel study of the ΔA model, based on an approach similar to ours, was carried out by Seifert, Berndl and Lipowsky [27]. It turns out that the ΔA model has the same catalog of stationary shapes as the SC model, but it predicts shape-transition sequences distinctly different from those of the SC model. It will be useful to give here a brief account of the main results of Seifert *et al.*

The ΔA model assumes that the equilibrium shape of a fluid lipid-bilayer vesicle minimizes the bending energy (2.30) under the constraints of fixed surface area A and volume V, and subject to an additional constraint on the area difference ΔA between the two leaves. ΔA is considered to have been fixed at the time of vesicle closure.[1] If M denotes the inte-

[1] Change in temperature certainly induces change in ΔA, which is the one of the factors driving shape transitions. We assume, however, that there is no lipid exchange between the two leaves on the timescale of experiments.

grated mean curvature, $M \equiv \oint dA(C_1 + C_2)/2$, ΔA can then be expressed as $\Delta A \approx 2DM$, as follows from (2.2). The constrained minimization for the ΔA model necessarily needs an additional Lagrange multiplier Q coupled to M. Thus, we introduce a new free energy

$$\Phi'[\Sigma', P', Q; S] = G[S] + \Sigma' A[S] - P'V[S] + QM[S] \,, \tag{4.1}$$

where $G[S]$ is given by (2.30), and where we have written Σ' and P' to emphasize that these are at this stage different parameters from Σ and P in (3.1). It is a natural outcome of the variational procedure that the Lagrange multipliers Σ', P' and Q are related to the first derivatives of the bending energy at the corresponding stationary point, i.e., for $G_m(A, V, M) \equiv G(A_m(\Sigma', P', Q), V_m(\Sigma', P', Q), M_m(\Sigma', P', Q))$,

$$\Sigma' = -\left.\frac{\partial G_m}{\partial A}\right|_{V,M}, \qquad P' = \left.\frac{\partial G_m}{\partial V}\right|_{A,M}, \qquad Q = -\left.\frac{\partial G_m}{\partial M}\right|_{A,V} . \tag{4.2}$$

An important observation can now be made: Q couples linearly to the integrated mean curvature, $M = \oint dA\,(C_1 + C_2)$. This coupling term has the same form as a cross term in the bending energy of the SC model. Thus, $\Phi'[\Sigma', P', Q; S]$ has exactly the same form as the free energy $\Phi[\Sigma, P; S]$ (3.1), provided that

$$\Sigma' = \Sigma + \kappa C_0^2/2, \quad P' = P, \quad \text{and} \quad Q = -2\kappa C_0 \,. \tag{4.3}$$

Thus, the Euler-Lagrange equations for the stationary shapes which follow from Eq. (4.1) are identical to those which follow from Eq. (3.1), except for a trivial change of notation, so (as noted in the previous paragraph) the two models share the same catalog of stationary shapes. Explicitly put, $S_n(\Sigma, P, C_0) = S_n(\Sigma', P', Q)$, provided that (4.3) is satisfied. In what follows, we will no longer make a distinction between (Σ',P',Q) and (Σ,P,C_0), and we shall in practice use the variables (Σ,P,C_0). Hence, we rewrite Eq. (4.2) as

$$-\left.\frac{\partial G_m}{\partial A}\right|_{V,M} = \Sigma + C_0^2/2, \qquad \left.\frac{\partial G_m}{\partial V}\right|_{A,M} = P, \qquad \left.\frac{\partial G_m}{\partial M}\right|_{A,V} = 2C_0 \,. \tag{4.4}$$

Although the two models predict the same catalog of stationary shapes, it does not follow that the phase diagrams of the two models are identical, since it is now the bending energy G (rather than the bending energy E) that must be used in selecting the overall

minimum-energy shape from among the stationary shapes $S_m(A, V, M)$. A phase diagram which summarizes the systematics of the vesicle shapes and shape transitions predicted by the ΔA model is illustrated in Fig. 4.1.

There are still three length scales in the problem, R_A, R_V, and M, even though the spontaneous curvature becomes irrelevant now (see Section 2.2). The bending energy G given in (2.30) is invariant under uniform dilation of the shape S by a factor λ, a fact which implies that

$$G(A, V, M) = G(\lambda^2 A, \lambda^3 V, \lambda M) . \tag{4.5}$$

It is, therefore, convenient to choose as variables for the phase diagram values of the volume and area difference scaled by the "area" length R_A,

$$v \equiv \frac{V}{4\pi R_A^3/3}, \quad \Delta a \equiv \frac{\Delta A}{8\pi R_A D} \equiv \frac{M}{4\pi R_A} . \tag{4.6}$$

Hence, $G(v, \Delta a) = G(A, V, M) = G(\lambda^2 A, \lambda^3 V, \lambda M)$ uniquely defines the phase diagram. Note that $0 \leq v \leq 1$. The point $v = 1$, $\Delta a = 1$ corresponds to spherical shapes and is a useful point of reference.

Fig. 4.1 shows the part of the full $(v, \Delta a)$ phase diagram in which budding and vesiculation occur. Starting above and to the left of the sphere point, we see a region of (prolate) elliptical shapes. Under increasing Δa and decreasing v, these shapes evolve via a transition which breaks up/down symmetry to "pear" shapes, which transform smoothly into narrow-necked shapes and finally reach vesiculation. Similarly, decreasing Δa leads through oblate ellipses via stomatocytes to internally vesiculated or invaginated shapes. Much of this is broadly similar to what we have seen in the SC model; however, two features associated with these shape transformations are crucially different: First, the driving mechanism for the budding is different. The constraint on the area difference produces a couple which tends to bend the surface in much the same way that a temperature-induced length difference bends a bimetallic strip. The larger ΔA gets, the more the vesicle surface tends to bend, in order to be convex outward towards the side with the larger area. Thus, large positive Δa (in our convention) favors small spherical buds on the outside (and small Δa favors small spherical buds on the inside) to extent allowed by the constraints on A and V.

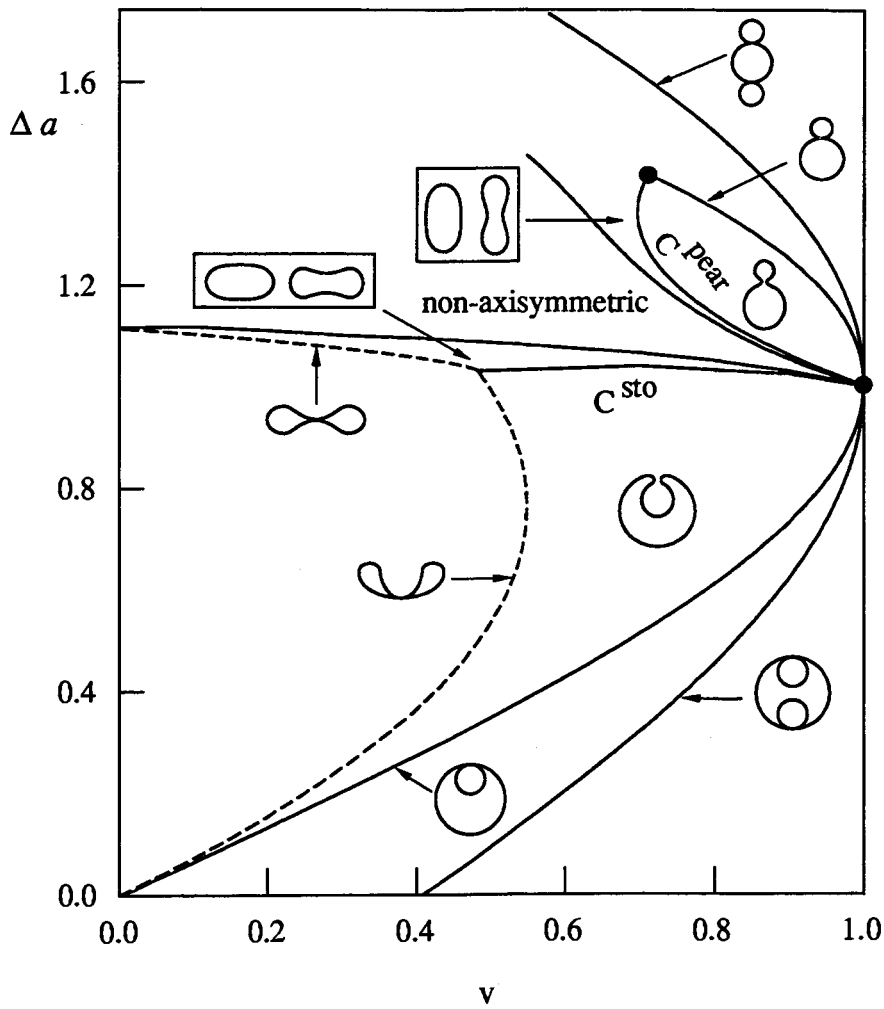


Figure 4.1: Phase diagram for the ΔA model (after Ref. [27]). C^{pear} and C^{sto} denote the lines of continuous transitions at which the up/down symmetry of the vesicle is broken. Beyond the dashed line additional interactions have to be taken into account in order to prevent self-intersection of the membrane.

A budded state conveniently accommodates this trend. Secondly, the budding[2] predicted by this model, whether it is outward or inward, is always *second order* (i.e., continuous) and proceeds via the pear-shaped intermediary, in contrast to the *first order* (i.e., discontinuous) budding characteristic of the SC model.

The bending energies for the stationary shapes of the SC and ΔA models are related by a transformation similar to a Legendre transformation. To see this, we rewrite the bending energy in the SC model as

$$E[A, V, C_0; S] = G[A, V, M; S] - 2\, C_0\, M + \frac{1}{2}\, C_0^2\, A\,, \tag{4.7}$$

where $G[A, V, M; S]$ is the bending energy Eq. (2.30) in the ΔA model. Imagine that we perform the constrained minimization of $E[A, V, C_0; S]$ at given A, V, and fixed C_0 in two steps: First, we minimize the expression on the right-hand side of Eq. (4.7) with respect to those shape variations that keep A, V, and M constant. This leads to the stationary shapes of the ΔA model, with

$$E(A, V, C_0; M) = G_n(A, V, M) - 2\, C_0\, M + \frac{1}{2}\, C_0^2\, A\,. \tag{4.8}$$

$G_n(A, V, M)$ is the bending energy of a stationary shape which belongs to a branch n, and there may be many branches for given (A, V, M). We may now make the energy function $E(A, V, C_0; M)$ stationary against arbitrary variations of M to determine its extrema,

$$\frac{\partial E(A, V, C_0; M^*)}{\partial M^*} = \frac{G_n(A, V, M^*)}{\partial M^*} - 2\, C_0 = 0\,, \tag{4.9}$$

which determines M^* as a function of A, V, and C_0. The bending energy functional E at its stationary points is, therefore, given by the following transform of $G_n(A, V, M^*)$, which is very similar to a Legendre transform in spirit:

$$E_n(A, V, C_0) = G_n(A, V, M^*) - 2\, C_0\, M^* + \frac{1}{2}\, C_0^2\, A\,. \tag{4.10}$$

It is sometimes more convenient to write everything in dimensionless form. In this form,

[2]Note that, in this continuous route to the narrow-necked, budded state the only "thermodynamic" singularity is at the up/down symmetry-breaking transition. The further evolution of the pear to the narrow neck is entirely smooth.

we define the "reduced spontaneous curvature," $c_0 \equiv C_0 R_A$, and we find [3]

$$E_n(v, c_0) = G_n(v, m^*) - 2\, c_0\, m^* + 2\pi\, c_0^2\,,$$
$$\frac{\partial G_n(v, m^*)}{\partial m^*} = 2\, c_0\,. \tag{4.11}$$

This result will be useful when we make a connection between the ADE model and its limiting cases, i.e., the SC and ΔA models, in Section 4.3.

Which model describes correctly the physics of a fluid lipid-bilayer vesicle? The answer can only come from a careful comparison between the theoretical predictions and the available experimental evidence, such as that summarized in Chapter 1. Since both the SC and ΔA models are capable of producing the same catalog of stationary shapes,[4] they cannot be distinguished by looking just at the shapes of single vesicles. It is necessary to consider the systematics of the sequence of shape transitions [29]. It turns out that the characteristics of the vesicle shape transitions revealed by the experiments is more complex than the predictions of either of the two models.

We shall concentrate in the following on outward budding and vesiculation, exclusively, since they are the most prominent shape transformations seen experimentally and observations are, correspondingly, more extensive and systematic [9, 20] than for other shape changes. The temperature trajectories B and C described in Section 1.2 show two typical scenarios for budding and vesiculation. The thermally induced *discontinuous* budding involved in trajectory C is similar to the first-order transition from a prolate ellipsoid to a budded state in the SC model, a fact which makes this model attractive. However, the SC model is really unsatisfactory: It postulates an asymmetry represented by C_0. According to the SC picture, the asymmetry which causes $C_0 \neq 0$ is chemical in origin, resulting, for example, from a difference in composition between the interior and exterior leaves of the

[3] The unit length chosen here, as in Eq. (4.6), is the "area" length R_A. This choice of the dimensionless variables, i.e., v and c_0, is consistent with the normalization of length scale, $R_A = 1$. This choice is an alternative to our previous one, Δ and R_V, as made for the calculations in Chapter 3. c_0 is therefore equivalent to the "area" length in Chapter 3, which is scaled by C_0 and equals to $R_V\,(1+\Delta)^{1/2}$. Thus, $c_0 = R_V\,(1+\Delta)^{1/2}$, and $v = 1/(1+\Delta)^{3/2}$.

[4] As far as predicting equilibrium shapes is concerned, both models are quite successful. All the stationary shapes obtained in theoretical calculations have been observed in experiments.

bilayer or from different screening properties of the interior and exterior aqueous environments. If this picture were valid, all vesicles prepared in the same way should have the same C_0, and the direction (outward or inward) of vesiculation should be the same for all vesicles in a given preparation. This prediction is simply not consistent with the observed diversity of shape sequences, as outlined in Chapter 1. Another problem is that, if the *continuous* symmetry-breaking transition that takes a prolate to a pear shape (see Sequence B in Section 1.2) is, indeed, an equilibrium transition (referred to as *continuous budding* in later text), the SC model cannot provide an explanation for it. The ΔA model seems more promising from this perspective: Continuous symmetry breaking is a characteristic feature of this model. Furthermore, the diversity of shapes observed in a given preparation can be explained within the context of the ΔA model by the hypothesis that different vesicles in the preparation have different values of ΔA as a result of the (unknown) closure process, when the vesicle first formed. Indeed, a variety of observed thermal trajectories have all been explained by postulating a small asymmetry in the thermal expansion coefficient of the two leaves of the bilayer [23]. Despite this impressive success however, the ΔA model does not give a full picture, since it cannot explain either the discontinuous budding in Sequence C (Section 1.2) or the discontinuous transition observed in Sequence B between the weakly asymmetric pear and the narrow-necked or completely vesiculated shape. We are, therefore, inclined to believe that there are physical variables and effects which have not yet been properly controlled or incorporated into the two models.

Moreover, it has also been shown that pretreatment of a vesicle before observation affects its transition sequence. For example, precooling a vesicle for several hours always induces the outward budding, whereas vesicles without the same treatment show other transition sequences, such as a transition to the stomatocyte followed by inverse vesiculation [20]. This suggests that the history of the vesicle formation and the treatment prior to the observation must be considered as variables relevant to shape transitions.

The ADE model seems to have the potential to solve some but not all of these problems. Since it contains an additional coupling $\bar{\kappa}$ (see Eq. (2.1)), it provides another control parameter which may be important in explaining the diversity of different observed behaviors.

Furthermore, it interpolates between the SC and ΔA models (Section 2.2) and, so, may be expected to encompass both discontinuous budding directly from the up/down-symmetric state (as in the SC model) and continuous up/down symmetry breaking (as in the ΔA model). It is the purpose of this chapter to explore how the presence of the non-local bending energy is reflected in equilibrium shapes and shape transitions of vesicles, and to discuss whether this model is able to reconcile theory with experiments.

The technical framework within which the systematic analysis of the ADE model should be performed is, in principle, the same as for the two conventional models. The equilibrium shape is determined by minimizing the bending energy (2.1) for given surface area A, volume V, and for fixed spontaneous curvature C_0 and optimal area difference ΔA_0 between the two monolayers. For the sake of convenience, we recall the ADE bending energy (2.1),

$$W[\alpha, A, V, C_0, \Delta A_0; S] \equiv \frac{1}{2}\oint dA[C_1(\mathbf{r}) + C_2(\mathbf{r}) - C_0]^2 + \frac{\alpha}{2}\frac{\pi}{AD^2}(\Delta A - \Delta A_0)^2 \,, \quad (4.12)$$

where κ is set to 1. One might expect that the additional non-local bending energy would complicate the analysis. Fortunately, the same catalog of stationary shapes seen in the SC and ΔA models emerges again from the ADE model: Suppose that $S_n(\Sigma, P, \bar{C}_0)$[5] is a stationary shape in the ΔA model, corresponding to v and $m \equiv 4\pi\Delta a$. This same shape will also be a stationary shape in the ADE model, for the same v and some fixed α, c_0 and $m_{0,n} \equiv \Delta A_0/(2DR_A)$, if it is specified that[6]

$$m_{0,n} = m + (G_n' - 2c_0)/\alpha \,, \quad (4.13)$$

where $G_n' \equiv \left.\frac{\partial G_n}{\partial m}\right|_v = 2\,\bar{c}_0$ denotes the derivative of the bending energy $G_n(v, m)$ at fixed v along the branch of stationary shapes to which $S_n(\Sigma, P, \bar{C}_0)$ belongs. Making use of Eq. (4.13) to evaluate the bending energy $W(\alpha, v, c_0, m_0)$ defined in Eq. (4.12), we obtain the bending energy for stationary shapes in the ADE model as follows:

$$W_n(\alpha, v, c_0, m_0) = G_n(v, m) - 2c_0 m + 2\pi c_0^2 + \frac{\alpha}{2}(m - m_0)^2. \quad (4.14)$$

[5]We have changed notation slightly, using $\bar{C}_0$ to denote the Lagrange multiplier coupled to the integrated mean curvature M, which was represented previously by C_0 in Eq. (4.4). The purpose is to distinguish it from the physical parameter C_0 (the spontaneous curvature).

[6]This relationship is established by the same "two-step variation" argument used in deriving the Legendre transformation (4.9) and (4.10) between the SC model and the ΔA model.

The use of the dimensionless variables v, c_0, and m_0 eliminates the redundancy associated with the invariance of the bending energy under dilations (see Eq. (4.5)). Combined with α, the variables v, c_0, and m_0 span a complete four-dimensional phase space, in which the systematics of equilibrium shapes and shape transformations coming out of our analysis can be illustrated.

In view of the connection with previous work on the SC and ΔA models, the detailed investigation is rather straightforward; however, its outcome is non-trivial. In particular, for fixed values of the elastic constants with $0 < \bar{\kappa}/\kappa \equiv \alpha < \infty$, we find budding sequences which are both continuous (second order), as in the ΔA model, and discontinuous (first order), as in the SC model. Which sequence a particular vesicle will follow depends on the value of its initial, relaxed area difference ΔA_0. It will not be difficult to relate ΔA_0 to the treatment vesicles receive prior to shape-transformation experiments, as we shall discuss in Section 4.4.2.

The remainder of this chapter is organized as follows: In Section 4.2, we derive the phase diagram of the ADE model for the special case in which $C_0 = 0$. In particular, we show how the characteristics of budding depend on the value of α. In Section 4.3, we generalize the results obtained in Section 4.2 to the case where the spontaneous curvature C_0 is non-zero. We show that the effect of a non-zero spontaneous curvature leads to a nearly trivial modification of the $C_0 = 0$ phase diagram. We also discuss how the phase diagram in the ADE model interpolates between those in the SC and ΔA models. In Section 4.4, we discuss the relation of the predictions of the ADE model to the available experiments: We calculate temperature trajectories in phase space, which show the shape evolution of a vesicle in a temperature-driven experiment. Estimates for the model parameters α and ΔA_0 for the experimentally studied systems SOPC and DMPC suggest that it is the first-order scenario which should be expected. In fact, apparent second-order budding is frequently observed. We discuss in Section 4.4.2 the current status of attempts to reconcile theory with experiments.

4.2 Budding and vesiculation I: Zero spontaneous curvature

We discuss first behavior when the spontaneous curvature C_0 is zero. This case deserves special attention, since vesicles formed by a perfectly symmetric lipid-bilayer fall into this category (see Chapter 2). Also, all generic features of the model are already present in this case and are best illustrated here. Generalization to the case where $C_0 \neq 0$ is straightforward, as we shall demonstrate in Section 4.3.

The budding and vesiculation we shall focus on in this chapter involve two free-energy branches of axisymmetric stationary shapes: The *symmetric branch* consists of shapes having mirror (or up/down) symmetry with respect to their equatorial plane and resembling prolate ellipses or dumbbells. The *asymmetric branch,* in which the mirror symmetry is broken, is composed of the so-called "pear-like" (with wide neck) or "budded" (with narrow neck) shapes. This branch ends with a fully vesiculated shape, where the neck radius has formally shrunk to zero.

4.2.1 The phase diagram at $\alpha = 4$: Discontinuous budding versus continuous budding

For $c_0 = 0$, the phase diagram is still three-dimensional, corresponding to parameters (α, v, m_0), where $\alpha \equiv \bar{\kappa}/\kappa$, $v \equiv V/(4\pi R_A^3/3)$, and $m_0 \equiv \Delta A_0/(2\,D R_A)$. A typical, two-dimensional fixed-α section is shown for $\alpha = 4$ in Fig. 4.2. Symmetric prolate ellipses and dumbbells are the lowest-energy shapes in the region below the two-part curve labelled[7] C^{pear} and D^{pear}. The lower bound of this region is delimited by a transition either to a non-axisymmetric ellipsoid or to an oblate shape, depending on α, v, and m_0. Since these transitions are not at the center of the problem being addressed in this chapter, we will not discuss them in any detail. At the line denoted C^{pear}, the symmetric shapes become locally unstable with respect to a mode that breaks the mirror symmetry and undergo a continuous transition to pear-like shapes which take over as the lowest-energy configurations

[7]We will denote all phase boundaries by superscripted capital letters: C stands for "continuous"; D, for "discontinuous"; M, for "metastable"; L, for "limiting"; etc.

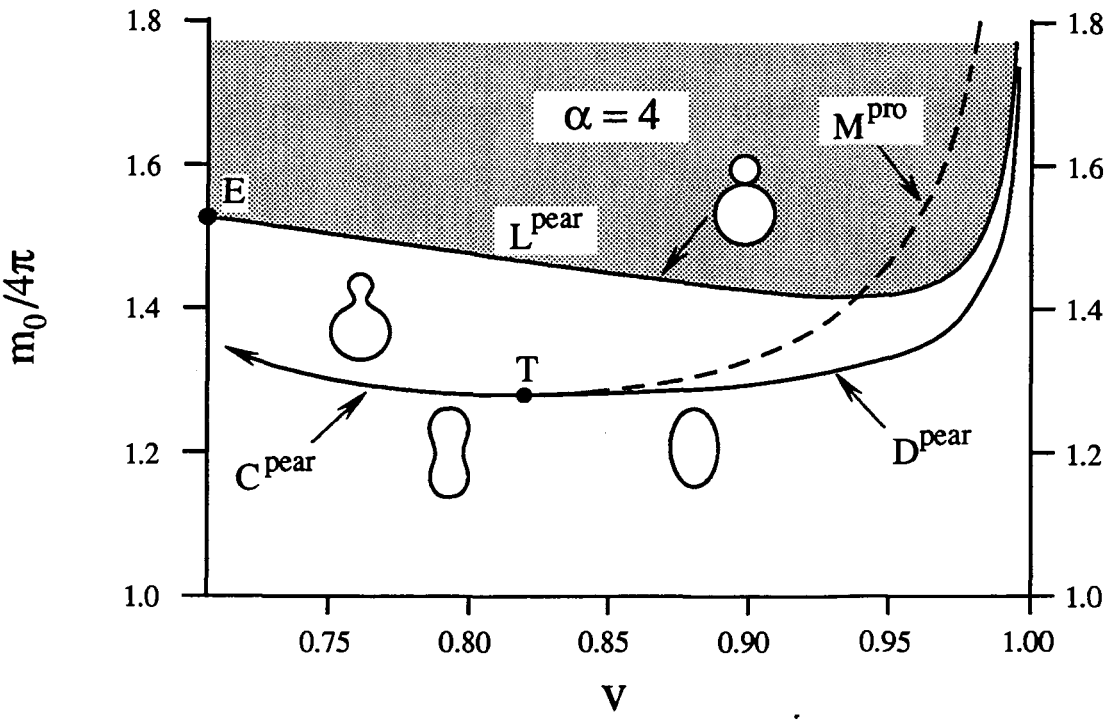


Figure 4.2: Phase diagram in the (v, m_0)-plane involving budding and vesiculation at fixed $\alpha = 4$. Symmetric and asymmetric shapes are separated by lines of continuous transitions, C^{pear}, or discontinuous transitions, D^{pear}. These lines of budding transitions meet at the tricritical point T. The line M^{pro} denotes the limit of metastability of the symmetric shapes. The line L^{pear} denotes the vesiculation line, where the neck size has shrunk to zero. It ends at the point E, where the limiting shape consists of two spheres of the same size. In the shaded region above this line, multiplets involving several buds become relevant (see Ref. [28]).

in the region above C^{pear}. The phase boundary D^{pear} at larger reduced volumes marks the onset of an instability of a different character: Above D^{pear} and below M^{pro}, the symmetric shapes remain locally stable but are unstable globally, since the asymmetric shapes have lower overall bending energy. The transition is therefore discontinuous or "first-order". The continuous phase boundary C^{pear} and the discontinuous phase boundary D^{pear} meet at the special point T with coordinates $(v_T(\alpha), m_{0,T}(\alpha))$, which we call a "tricritical point" by analogy with thermodynamics. Bounded below by C^{pear} and D^{pear} and above by L^{pear} is the single-phase region of asymmetric shapes. The neck radii go to zero continuously as L^{pear} is approached.

In constructing the phase diagram Fig. 4.2, we exploited the relationship between the ΔA model and the ADE model established in Eqs. (4.13) and (4.14). In particular, the bending energy W for each branch of stationary shapes may be expressed as

$$W_n(\alpha, v, m_0(m)) = G_n(v, m) + \frac{\alpha}{2}[m - m_0(m)]^2 \,, \tag{4.15}$$

while its first and second derivatives are related to those of the bending energy, $G_n(v, m)$, according to

$$\left.\frac{\partial W_n}{\partial m_0}\right|_{\alpha,v} = \left.\frac{\partial G_n}{\partial m}\right|_v \,, \tag{4.16}$$

and

$$\left.\frac{\partial^2 W_n}{\partial m_0^2}\right|_{\alpha,v} = \frac{\alpha G_n''(v, m)}{\alpha + G_n''(v, m)} \,, \tag{4.17}$$

where

$$G_n''(v, m) \equiv \left.\frac{\partial^2 G_n}{\partial m^2}\right|_v \,. \tag{4.18}$$

Note that $m \equiv 4\pi\Delta a$ (Eq. (4.6)). The details of the mapping, based on Eqs. (4.15)–(4.17), are illustrated in Fig. 4.3. The two upper plots in this figure show the generic structure of the bending energy G in the ΔA model as a function of m for fixed v, both for the symmetric branch and for the asymmetric branch [27]. At the bifurcation point, labeled by C in the figure, the energy curves of the two branches have the same first derivative. This property is preserved as m is mapped onto m_0 according to (4.13) and $G_{s,a}(v, m)$ are transformed to $W_{s,a}(\alpha, v, m_0)$ by (4.15). It is clear, therefore, that, in the neighborhood of the bifurcation point, the difference ΔW between the bending energies of the two branches depends upon the second derivatives only, i.e.,

$$\begin{aligned} \Delta W &\equiv W_s(\alpha, v, m_0) - W_a(\alpha, v, m_0) \\ &= \frac{1}{2}\left[W_s''(\alpha, v, m_{0,\mathrm{C}}) - W_a''(\alpha, v, m_{0,\mathrm{C}})\right](m_0 - m_{0,\mathrm{C}})^2. \end{aligned} \tag{4.19}$$

We can, thus, read off the local energetics from Eq. (4.19) near the bifurcation point $m_{0,\mathrm{C}} = m_\mathrm{C} + G'_{s,a}(v, m_\mathrm{C})/\alpha$, once we know the second derivatives of W. Eq. (4.17) enables us to acquire this information from the second derivatives of the bending energy G_n. We find, through direct numerical analysis, that the following properties hold generically for $0.72 < v < 1$:

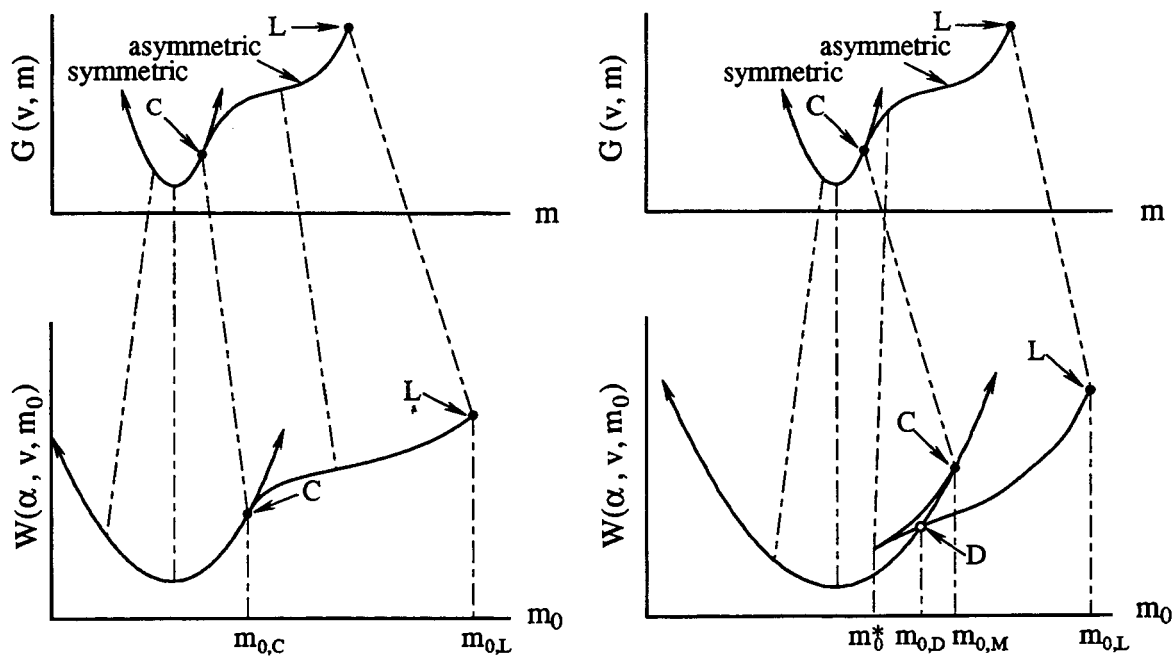


Figure 4.3: Schematic representation of the mapping between the ΔA and the ADE models, which involves the energy G as a function of m for the ΔA model and the energy W as a function of m_0 for the ADE model, respectively. L denotes the vesiculated shape where the neck size is zero and C denotes the bifurcation point. For large α, as displayed in (a), the topology of the bifurcation is preserved in the ADE model and the budding transition is still continuous. For small α, as in (b), the symmetric branch develops a wing structure in the ADE model, thus rendering the budding transition discontinuous.

(i) $G_s''(v,m) > 0$ for all values of m accessible to the symmetric branch;

(ii) For the asymmetric branch the situation is entirely different: $G_a''(v,m)$ starts being *negative* at the bifurcation point and then increases monotonically, reaching a positive value at the vesiculation point;

(iii) The magnitude of $G_a''(v,m_C)$ increases monotonically in v and eventually diverges in the limit $v \to 1$.

Finally, it is useful to know the behavior of the mapping (4.13) near the bifurcation point: Differentiating (4.13) with respect to m leads to the relation,

$$\frac{\partial m_{0,n}}{\partial m} = 1 + \frac{1}{\alpha} G_n''(v,m) \,. \tag{4.20}$$

At the bifurcation point, the relation is, explicitly,

$$\left.\frac{\partial m_{0,s}}{\partial m}\right|_{m=m_C} = 1+\frac{1}{\alpha} G_s''(v,m_C) \,, \qquad \text{and} \qquad \left.\frac{\partial m_{0,a}}{\partial m}\right|_{m=m_C} = 1+\frac{1}{\alpha} G_a''(v,m_C) \,. \tag{4.21}$$

These properties allow us to distinguish two different scenarios for the bifurcation in the ADE model, corresponding to the continuous transitions C^{pear} and the discontinuous transitions D^{pear}, respectively:

(i) For $v < v_T(\alpha)$ (with $v_T(\alpha)$ determined by $G_a''(v_T,m_C) = -\alpha$), the mapping (4.13) is monotonic for both the symmetric and the asymmetric branches, and ΔW is positive, as follows from Eqs. (4.17) and (4.19). Consequently, the local topology of the energy diagram of the ΔA model is well preserved,[8] as sketched in Fig. 4.3(a), and the symmetry-breaking transition remains continuous, as it is in the ΔA model. The corresponding phase boundary C^{pear} is simply given by

$$m_{0,C}(\alpha,v) = m_C(v) + \frac{1}{\alpha} G'_{s,a}(v, m_C(v)) \,. \tag{4.22}$$

(ii) For $v > v_T(\alpha)$, the function $m_0(m)$ defined by (4.13) becomes non-monotonic for the asymmetric branch, while it remains monotonic for the symmetric branch. In addition, for $m_0 \lesseqgtr m_{0,C}$, $\Delta W < 0$. The symmetric branch and the asymmetric branch thus interchange their relative positions locally, leading to a different global topology manifested by the

[8] The global energetics is not available in this kind of local analysis. In order to determine it, we performed the one-to-one mapping using Eq. (4.15) for the whole branches.

wing structure in the energy diagram shown in Fig. 4.3(b), where the cusp on the left hand side of the wing is determined[9] by $G_a''(v, m^*) = -\alpha$. The corresponding budding transition is, therefore, discontinuous and is represented by D^{pear}. The precise location of the first-order boundary has to be computed numerically. We performed such computations for $v < 0.90$. For larger values of the reduced volume v, numerical determination of the asymmetric branch becomes more cumbersome, due to the same kind of numerical instability of the differential shape equations that we have pointed out in Chapter 3 and Appendix B. However, approximating the very small-necked shapes by the vesiculated shapes, one can obtain a good approximation to D^{pear}.[10]

Related to the discontinuous transition is the issue of metastability. In particular, both symmetric prolate (or dumbbell) and asymmetric pear (or narrow necked) shape are locally stable at the transition D^{pear}. Above D^{pear}, the symmetric shapes remain locally stable until they eventually develop a local instability with respect to asymmetric perturbation at the (spinodal) line, denoted by M^{pro}, which is just the continuation of the line described by Eq. (4.22) beyond the tricritical point. Similarly, the asymmetric shape remains a local minimum of the bending energy below D^{pear}, only becoming locally unstable with respect to the symmetric disturbance at another spinodal point (M^{pear}). The location of M^{pear} is given by

$$m^*_{0,M}(\alpha, v) = m^*(v) + \frac{1}{\alpha} G_a'(v, m^*) \,, \tag{4.23}$$

where m^* is determined by

$$G_a''(v, m^*) = -\alpha \,. \tag{4.24}$$

For a given v, this spinodal point corresponds to the cusp in the wing structure (see Fig. 4.3), and must be found numerically.

The tricritical point T at which C^{pear} and D^{pear} meet has coordinates specified by

[9] Such a wing structure is analogous to the Landau-theory picture of a first-order transition, in which the wing structure eventually disappears at a critical point.

[10] As v approaches 1, the neck size of the budded shape at the first-order transition point becomes smaller and smaller. Since the bending energy W of a necked shape, when the neck is small, is very close to that of the corresponding (i.e., same area, volume and m_0) vesiculated shape, we may use the bending energy of the vesiculated sheet, $W_L(\alpha, v, m_0) = 2 \times 8\pi + (\alpha/2)[m_0 - m_L(v)]^2$, as a good approximation to that of the narrow-necked shapes.

$(v_T(\alpha), m_{0,T}(\alpha))$ where v_T satisfies the condition,

$$G_a''(v_T, m_C) = -\alpha \,, \tag{4.25}$$

and $m_{0,T}$ is given through the relationship (4.13).

The vesiculation boundary L^{pear}, corresponding to shapes having two coexisting spheres of different radii connected by an infinitesimal neck, is approached continuously by the pear-shaped vesicles and defines the upper limit of the "pear" region. This boundary may be parametrized in terms of the area difference Δa (Eq. (4.6)). In this form it is described by the following two equations:

$$\begin{cases} v_L = 1 - \frac{3}{2}(\Delta a - 1)^2 - \frac{1}{2}(\Delta a - 1)^3 \\ \\ m_{0,L} = 4\pi\Delta a + \frac{4}{\alpha}\,\frac{\Delta a}{(\Delta a^2 - 1)} \,. \end{cases} \tag{4.26}$$

With this we complete the derivation of the budding transition. The region above and to the left of the limiting line L^{pear} has not yet been seriously explored for the ADE model. For $0 < \alpha < \infty$ (i.e., away from the limiting cases), we expect, on the basis of our work on the SC model, a narrow strip where the lowest-energy configuration is a two-sphere multiplet, followed perhaps by a prolate (ellipse)-plus-sphere multiplet. Further, complex multiplets presumably form as ΔA_0 (or m_0) increases.

In the phase diagram Fig. 4.2 and throughout the above analysis, we have chosen reduced volume, $v > 0.72$. This choice is made in order to avoid a complete discussion of the neighborhood of the special point E ($v = \sqrt{2}/2, m_0 = 4\sqrt{2}(\pi + 1/\alpha)$), where the vesiculation line L^{pear} ends with a vesiculated shape consisting of two spheres of equal radii. This point is a limit point for various branches of shapes which approach E from $v < \sqrt{2}/2$. Due to the presence of these additional branches, a detailed analysis can be expected to be difficult. For most recent experiments, however, this point seems to be unimportant, since observed vesiculations have typically involved a bud of much smaller size than the mother vesicle, corresponding to a reduced volume $v \lesssim 1$.

Finally, we comment on the asymptotic behavior of the various phase boundaries shown

in Fig. 4.2 as $v \to 1$. We recapitulate here the asymptotic behavior of the vesiculation boundary L^{pear}, the discontinuous budding phase boundary D^{pear}, and the instability line M^{pro}. A detailed discussion is presented in Appendix E. The asymptotic behavior of the vesiculation boundary can be calculated as a series expansion in terms of $(1-v)^{1/2}$. The result is

$$m_{0,L}(\alpha, v) = m_L(v) + (2/\alpha)\bar{c}_{0,L}(v), \tag{4.27}$$

where, to the first few orders,

$$m_L(v) = 4\pi + 4\pi\sqrt{2/3}(1-v)^{1/2} + O[(1-v)], \tag{4.28}$$

and

$$\bar{c}_{0,L}(v) = \frac{\sqrt{3/2}}{(1-v)^{1/2}} + \frac{2}{3} - \sqrt{\frac{1}{6}} \times \frac{7}{12}(1-v)^{1/2} + O[(1-v)]. \tag{4.29}$$

The asymptotics of the boundary D^{pear} is given in leading order by

$$m_{0,D}(\alpha, v) = (2/\alpha)\frac{\sqrt{3/2}}{(1-v)^{1/2}} + O(1), \tag{4.30}$$

which is the same as the leading-order term in the expression for the vesiculation boundary. It is more difficult to deal with the asymptotics of the instability line M^{pro}. As sketched in Appendix E, we have reasons to believe that $m_{0,M}$ has also the leading divergence of the form $const./(1-v)^{1/2}$, where $const. > (2\sqrt{3/2})/\alpha)$.

4.2.2 Variation of the phase diagram with α: The SC and ΔA models as limiting cases.

We describe in this subsection a complete phase diagram for budding and vesiculation in the ADE model, which is illustrated in Fig. 4.4. The three-dimensional features of the phase diagram are illustrated by several sections corresponding to different values of α, for which the calculation discussed in the previous subsection is repeated. We discuss separately the situation for values of α smaller and larger than $\alpha = 4$.

(i) As the value of the parameter α decreases, the tricritical point moves towards smaller values of the reduced volume, i.e., $v_T(\alpha)$ decreases monotonically with decrease of α. It reaches the value 0.72, when α decreases to $\alpha_1 \simeq 1.2$. The issue of whether the tricritical

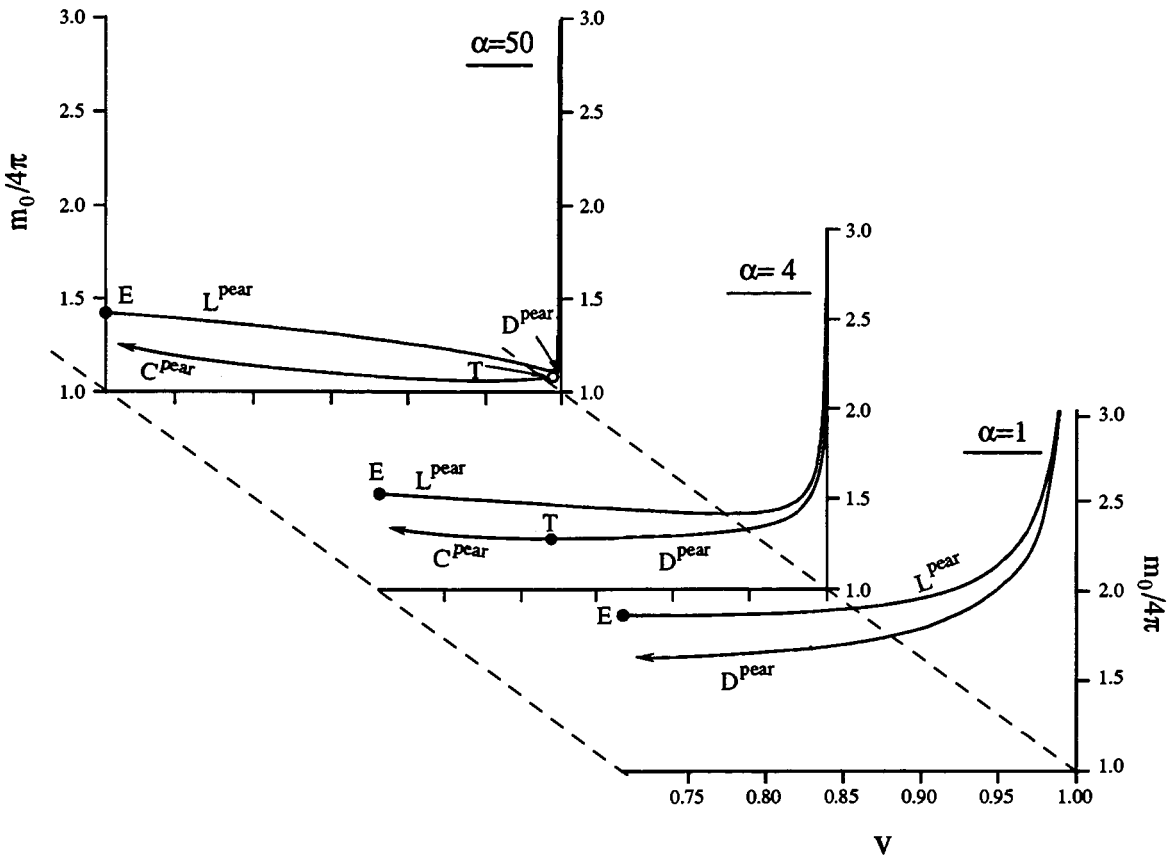


Figure 4.4: Phase diagram for the ADE model for various values of α. For $\alpha = 1$, the budding transition is always discontinuous (D^{pear}). For $\alpha = 50$, the budding is continuous (C^{pear}), except very close to $v = 1$. We can only provide an upper bound (in v and m_0) to the tricritical point denoted by the open circle.

point T persists or disappears, and how it persists or disappears, as α decreases from α_1, has not been explored thoroughly: Close to the special point E (the endpoint of the vesiculation boundary, at which the vesiculated shape involves two spheres of the same radii), many branches may participate in the competition for lowest-energy shape. The interplay of those branches needs to be understood, in order to resolve the further development of the tricritical point. It is clear however, that, when $\alpha < \alpha_1$, the budding transition is exclusively discontinuous, for $0.72 < v < 1$. As $\alpha \to 0$, we recover the results reported in Ref. [27] for the SC model with $C_0 = 0$. This limit is approached as follows: The discontinuous phase

boundary D^{pear} and the vesiculation boundary L^{pear} move towards infinite m_0 as

$$m_{0,D}(\alpha, v) = \frac{2\bar{c}_{0,D}(v)}{\alpha} + O(1), \tag{4.31}$$

and

$$m_{0,L}(\alpha, v) = \frac{2\bar{c}_{0,L}(v)}{\alpha} + m_L(v), \tag{4.32}$$

respectively. $\bar{c}_{0,D}(v)$ is the value of the reduced spontaneous curvature at the discontinuous transition point in the SC model. $\bar{c}_{0,L}(v)$ and $m_L(v)$ denote the values of those parameters at the vesiculation point. Consequently, at $\alpha = 0$, prolates (dumbbells) are the only equilibrium shapes for $v > 0.72$, irrespective of the value of m_0.

(ii) As the value of α increases from $\alpha = 4$, the tricritical point moves towards $v = 1$, and the budding transition is continuous throughout almost the whole range $1 > v > 0.72$. The budding and the vesiculation phase boundaries both approach those in the ΔA model, which are their proper limits as $\alpha \to \infty$, as indicated by Eqs. (4.22) and (4.27). However, the approach to the ΔA model is singular for $v \to 1$. For the vesiculation boundary, we can rewrite Eq. (4.27) using Eqs. (4.28) and (4.29),

$$m_{0,L}(\alpha, v) = 4\pi + (1-v)^{1/2}\left\{4\pi\sqrt{2/3} + O[(1-v)^{1/2}] + \frac{1}{\alpha(1-v)}\left[2\sqrt{3/2} + O[(1-v)^{1/2}]\right]\right\}. \tag{4.33}$$

It follows that, as $\alpha \to \infty$ at fixed v (but very close to 1), $m_{0,L}$ approaches 4π. On the other hand, $m_{0,L}$ diverges as $1/(1-v)^{1/2}$ when $v \to 1$ at fixed α. Using results and arguments from Appendix E, we find for the continuous phase boundary C^{pear},

$$\begin{aligned} m_{0,C}(\alpha, v) &= 4\pi + (4\pi/3)(1-v) + O[(1-v)^{3/2}] + \frac{1}{\alpha}\left[\frac{const.}{(1-v)^{1/2}} + O(1)\right] \\ &\approx 4\pi + (1-v)\left[4\pi/3 + \frac{const.}{\alpha(1-v)^{3/2}}\right], \end{aligned} \tag{4.34}$$

which ends at the tricritical point. We were not able to perform a rigorous calculation for the behavior of the tricritical point in the limit $\alpha \to \infty$. Our best guess is that $1 - v_T(\alpha) \sim 1/\alpha$ and $m_{0,T}(\alpha) \approx 4\pi + const./\alpha^{1/2}$.

4.3 Budding and vesiculation II: Nonzero values of the spontaneous curvature

Nonzero values of the spontaneous curvature C_0 result from the fact that a lipid bilayer may have a tendency to curve one way or the other due either to intrinsic chemical asymmetry of the two constituent monolayers or to a chemical asymmetry of the adjacent (interior and exterior) fluids (see Section 2.1). Both these effects are likely to occur in many systems of lipid-bilayer vesicles. We devote this section to a discussion of the budding transition in the case where C_0 is nonzero.

The generalization of our discussion in the previous sections to this case is, in fact, technically straightforward. By defining a new variable,

$$\bar{m}_0 \equiv m_0 + \frac{2}{\alpha} c_0 \,, \tag{4.35}$$

we can express the bending energy (4.14) as

$$W_n(\alpha, v, c_0, \bar{m}_0) = G_n(v, m) + \frac{1}{2}\alpha(\bar{m}_0 - m)^2 + (2\pi + \frac{2}{\alpha})c_0^2 - 2c_0\bar{m}_0 \,, \tag{4.36}$$

where $\bar{m}_0$ and m are related by the mapping,

$$\bar{m}_0 = m + \frac{1}{\alpha} G_n'(v, m) \,. \tag{4.37}$$

For a fixed nonzero c_0, the functional dependence of the bending energy W_n on $\bar{m}_0$ differs from that in the case of $c_0 = 0$ only by a constant and a term linear in $\bar{m}_0$ (as shown by Eq. (4.36)), neither of which enters the second derivative of W_n in Eq. (4.36). Thus, our previous discussion both of the preservation of the topology of the energy curves by the mapping and of the location of the tricritical points hold true again here. Explicitly, the location of the continuous symmetry-breaking phase boundary $m_{0,C}(c_0)$ is, for fixed c_0,

$$m_{0,C}(c_0, v) = m_{0,C}(c_0 = 0, v) - \frac{2c_0}{\alpha} = m_C(v) + \frac{1}{\alpha}\left[G_{s,a}'(v, m_C) - 2c_0\right] \,. \tag{4.38}$$

The discontinuous budding transition occurring when $v > v_T$ has to be determined from

$$W_s(\alpha, v, c_0, \bar{m}_0) = W_a(\alpha, v, c_0, \bar{m}_0) \,, \tag{4.39}$$

which simplifies to

$$G_s(v, m_s(\bar{m}_0)) + \frac{\alpha}{2}[\bar{m}_0 - m_s(\bar{m}_0)]^2 = G_a(v, m_a(\bar{m}_0)) + \frac{\alpha}{2}[\bar{m}_0 - m_a(\bar{m}_0)]^2. \qquad (4.40)$$

This condition is identical to the one which determines the discontinuous budding transition for the case $C_0 = 0$, only with m_0 now replaced by $\bar{m}_0$. The solution is, then,

$$\bar{m}_{0,\mathrm{D}}(v) \equiv m_{0,\mathrm{D}}(c_0, v) + 2c_0/\alpha = m_{0,\mathrm{D}}(c_0 = 0, v)\,, \qquad (4.41)$$

where $(v, m_{0,\mathrm{D}}(c_0 = 0, v))$ defines the discontinuous phase boundary $\mathrm{D}^{\mathrm{pear}}$ in Fig. 4.2.

These results show that the fixed-α section of the budding phase diagram for a nonzero value of the spontaneous curvature c_0 is identical to the $c_0 = 0$ phase diagram (see Fig. 4.2) except for a shift of m_0 by an amount $-(2/\alpha)c_0$.

The ADE model recovers the SC model and the ΔA model as its two limiting cases as $\kappa \to 0$ and $\kappa \to \infty$, just as for $C_0 = 0$. We illustrate this feature in Fig. 4.5, a three-dimensional phase diagram for budding and vesiculation in the parameter space (α, c_0, m_0) for a typical (fixed) value of the reduced volume ($v = 0.8$). In the limit $\alpha \to \infty$, which corresponds to the ΔA model, the continuous symmetry-breaking transition $\mathrm{C}^{\mathrm{pear}}$ and the final vesiculation $\mathrm{L}^{\mathrm{pear}}$ are represented asymptotically by the two straight lines at $m_0 = m_{\mathrm{C}}(v)$ and $m_0 = m_{\mathrm{L}}(v)$ [27], respectively, independent of c_0. As $\alpha \to 0$, the budding and vesiculation boundaries are again two straight lines, $c_0 = \bar{c}_{0,\mathrm{D}}(v)$ and $c_0 = \bar{c}_{0,\mathrm{L}}(v)$ [27, 28], independent of m_0. The tricritical line,

$$\begin{cases} m_{0,\mathrm{T}}(c_0, v) = m_{\mathrm{C}}(v) + \left[G'_{s,a}(v, m_{\mathrm{C}}) - 2c_0\right]/\alpha_{\mathrm{T}}(v) \\ \\ \alpha_{\mathrm{T}}(v) = -G''_a(v, m_{\mathrm{C}})\,, \end{cases} \qquad (4.42)$$

separates the continuous bifurcations, denoted by the sheet C, from the discontinuous budding, denoted by the sheet D. The asymptotic limits of each of the two sheets, as $\alpha \to \infty$ and $\alpha \to 0$, are two straight lines perpendicular to each other. The twist of the sheets C and D and the splitting of the sheet D along the line $\bar{c}_0 = \bar{c}_{0,\mathrm{D}}(v)$ may then be understood as a necessity for accommodating the peculiar geometry imposed by these two limiting cases.

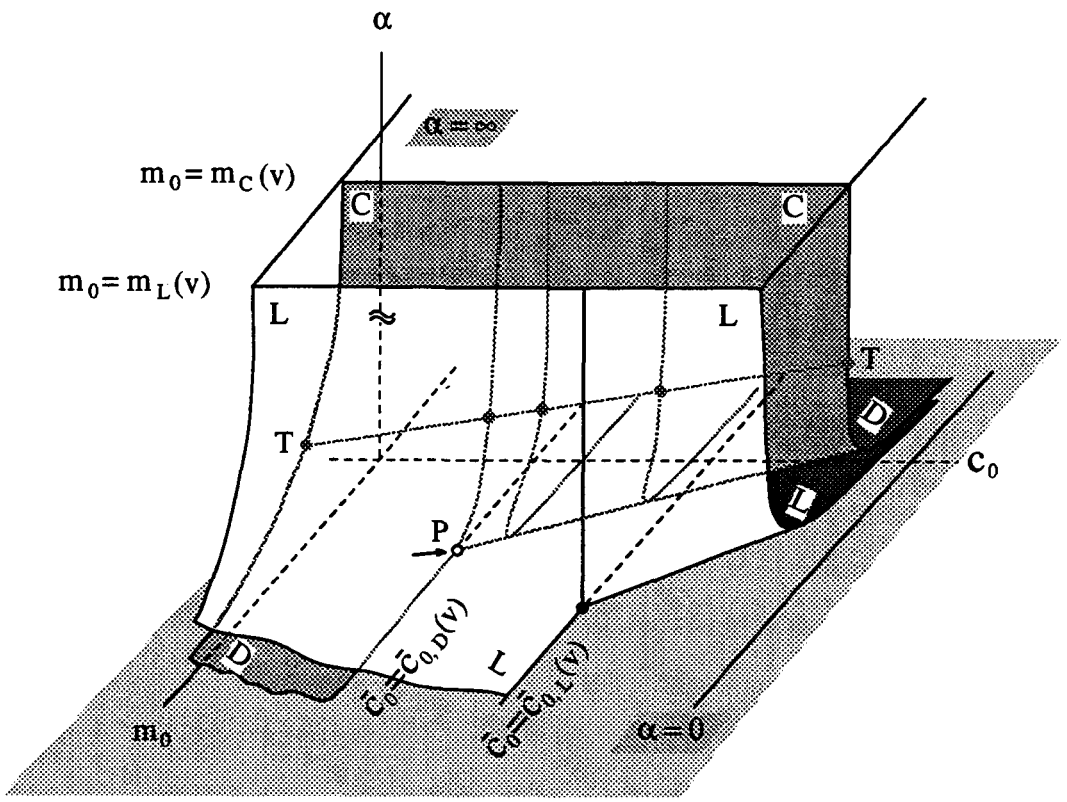


Figure 4.5: Three-dimensional (α, c_0, m_0) phase diagram for the ADE model at a typical (constant) reduced volume. Two transition surfaces are shown. The rear one (with darker shading) corresponds to budding and the front one corresponds to vesiculation. As $\alpha \to \infty$ (one of the limiting cases), the two sheets asymptotically approach their limits, represented by the two lines $m_0(v) = m_C(v)$ and $m_0(v) = m_L(v)$, respectively. These limit lines are independent of C_0, as in the ΔA model; when $\alpha \to 0$, the discontinuous budding at $\bar{c}_{0,D}$ and the final vesiculation at $\bar{c}_{0,L}$ predicted by the SC model are asymptotically recovered. T denotes the tricritical line, which separates the continuous transitions, represented by the sheet C, from the discontinuous budding transitions, represented by the sheet D. The budding sheet splits along $c_{0,D} = \bar{c}_{0,D}$ and the two parts develop in opposite directions. The splitting terminates at the point P with coordinates $(\alpha = 0, c_0 = \bar{c}_{0,D}, m_0 = (m_{s,D} + m_{a,D})/2)$, where $m_{s,D}$ and $m_{a,D}$ are the actual area differences of the symmetric and asymmetric shapes at the first-order transition point in the SC model. The vesiculation sheet has a similar topology, with a splitting along the line $c_0 = \bar{c}_{0,L}$. Its splitting terminates at $(\alpha = 0, c_0 = \bar{c}_{0,L}, m_0 = m_L)$.

4.4 Discussion: Relation to experiments

In Section 1.2, we described certain vesicle-shape transition sequences seen in the laboratory. In this Section, we wish to examine the question of to what extent the theoretical models we have presented are compatible with those observations. Most (but not all) of the experiments involve temperature-driven shape transformations. Thus, before beginning, it is important to establish (a) how the model parameters A, V, ΔA_0, α, and C_0 may be expected to vary with temperature and (b) how the trajectories through parameter space defined by $v(T)$, $m_0(T)$, etc., intersect the phase diagrams we have presented in the preceding Sections.

4.4.1 Temperature trajectories

Temperature increase leads to a thermal expansion of both the membrane area A of the vesicle and the volume V of its enclosed aqueous interior. We focus on the simple case where the vesicle is composed of two identical monolayers. It makes sense, therefore, to assume that both monolayers have the same thermal expansivity,[11]

$$\beta_A \equiv \frac{1}{A^{out}}\frac{dA^{out}}{dT} \equiv \frac{1}{A^{in}}\frac{dA^{in}}{dT} \,. \tag{4.43}$$

Consequently, the bilayer, effectively represented by the neutral surface (see Eq. (2.21)), has the same thermal expansion as the two monolayers,

$$\frac{1}{A}\frac{dA}{dT} = \beta_A \,. \tag{4.44}$$

The volume expansivity of the enclosed aqueous fluid (principally water) is typically appreciably smaller than the membrane area expansivity. Typical values are $\beta_V(\mathrm{H_2O}) \simeq 3 \times 10^{-4}/K$ and $\beta_A \simeq 4 \times 10^{-3}/K$ for SOPC and DMPC [22]. Thus, neglecting volume expansion, one obtains the temperature variation of the reduced volume v as

$$\frac{dv}{dT} = -(3\beta_A/2)v \,. \tag{4.45}$$

[11] Of course, one should not at all exclude the possibility that some kind of residual impurities within each of the constituent monolayers may create an asymmetry in the thermal expansivities, a fact which has been explored in Ref. [27] and takes on a special importance in the ΔA model.

The dependence of m_0 upon temperature is determined not only by the thermal behavior of the membrane area, but also by that of the bilayer thickness D. The temperature dependence of D has been measured for the fluid bilayers of several phospholipids, including POPC and DMPC, by using the NMR technique [56, 57]. In their fluid state ("liquid-crystalline" state, in the terminology used in the references), these phospholipid bilayers show a decrease in thickness as temperature is increased, indicating a negative thermal expansivity for the thickness,

$$\beta_D \equiv \frac{1}{D}\frac{dD}{dT} < 0 \,, \tag{4.46}$$

presumably due to entropic effects for the tails. In fact, the volume of a bilayer membrane, for a number of lipids, has been shown to change by only about 0.07% per degree ([57], [58]) and can, therefore, be effectively treated as constant over the temperature range explored in experiments. This suggests that $\beta_D = -\beta_A$ should be a good first approximation to the thermal behavior of lipid bilayers, if we assume $V_{\text{bilayer}} \simeq AD$ fixed. With these assumptions, we find a simple temperature dependence for m_0,

$$\frac{dm_0}{dT} = (\beta_A/2 - \beta_D)m_0 = (3\beta_A/2)m_0 \,. \tag{4.47}$$

Eliminating T, we arrive at a simple relation between v and m_0,

$$m_0(T) = m_0(T_0)\frac{v(T_0)}{v(T)} \equiv \frac{\hat{m}_0}{v(T)} \,. \tag{4.48}$$

Temperature trajectories in the (v, m_0) plane can, thus, be parametrized by the reduced equilibrium area difference $\hat{m}_0$ for the corresponding spherical shape. For the sake of simplicity, we shall assume α to be temperature independent (since there is, so far, no systematic information on a possible temperature dependence of this parameter).[12] Finally, for a first analysis, we take $C_0 = 0$. Under these conditions, Eq. (4.48) completely specifies the temperature trajectories. It will be useful in what follows to refer to the particular temperature trajectories which pass through the tricritical point T $(v_T(\alpha), m_{0,T}(\alpha))$ and the special endpoint E $(\sqrt{2}/2, \sqrt{2}(4\pi + 4/\alpha))$ (see Fig. 4.6). We shall, in what follows, call

[12] However, one does have a crude picture, based on the microscopic derivation presented in Chapter 2: α depends principally on the ratio of the monolayer thickness to the bilayer thickness. This ratio should not depend strongly on temperature T.

them the T and E trajectories, respectively. The initial (i.e., $v = 1$) values of the reduced area difference for these trajectories are

$$\hat{m}_{0,\mathrm{T}} \equiv m_{0,\mathrm{T}}(\alpha) v_{\mathrm{T}}(\alpha) ,$$

$$\hat{m}_{0,\mathrm{E}} \equiv 4\pi + 4/\alpha . \tag{4.49}$$

The relative positions of the special trajectories T and E defined above change with α and, thereby, define three regimes. The trajectories T and E divide the phase plane (v,m_0) into regions of different budding/vesiculation behavior:

(A) For an intermediate value of α both trajectories T and E exist with $m_{0,\mathrm{T}} < m_{0,\mathrm{E}}$, as shown in Fig. 4.6 for $\alpha = 4$. Different initial area differences lead to different characteristic thermally induced budding trajectories.

(i) *Discontinuous budding with vesiculation.* For $\hat{m}_0 > \hat{m}_{0,\mathrm{E}}$, increasing temperature causes the vesicle to follow a trajectory that crosses the discontinuous budding transition and finally leads to a smooth vesiculation. Note, however, that the elliptical branch remains locally stable at the discontinuous phase boundary $\mathrm{D}^{\mathrm{pear}}$ and does not develop local instability until reaching the spinodal line $\mathrm{M}^{\mathrm{pro}}$. Between $\mathrm{D}^{\mathrm{pear}}$ and $\mathrm{M}^{\mathrm{pro}}$ is a region of metastability/hysteresis. Thus, upon slow heating, a prolate ellipse is expected to pass smoothly through the first-order boundary $\mathrm{D}^{\mathrm{pear}}$ and to become unstable only somewhat before $\mathrm{M}^{\mathrm{pro}}$, when the metastable barrier has diminished to order $k_B T$, so thermal fluctuations can drive the shape over the barrier to a new minimum of the free energy. For large values of $\hat{m}_0$ the region of prolate metastability can extend well beyond the vesiculation line $\mathrm{L}^{\mathrm{pear}}$, so the eventual instability may well be into a fully vesiculated shape. In this way, it is possible that increasing-temperature trajectories may miss entirely the region of stable pears. Of course, $\mathrm{M}^{\mathrm{pro}}$ eventually meets the tricritical point, so, for smaller values of $\hat{m}_0$, the spinodal is still in the pear region. In general, the larger $\hat{m}_0$ is, the smaller is the temperature interval between $\mathrm{D}^{\mathrm{pear}}$ and $\mathrm{L}^{\mathrm{pear}}$, the smaller is the size of the vesiculated bud, and the more likely is the hysteretic effect which misses the pear.

(ii) *Discontinuous budding trajectory without vesiculation.* For $\hat{m}_{0,\mathrm{T}} < \hat{m}_0 < \hat{m}_{0,\mathrm{E}}$, the spherical vesicle, upon heating, exhibits again a discontinuous symmetry breaking. With

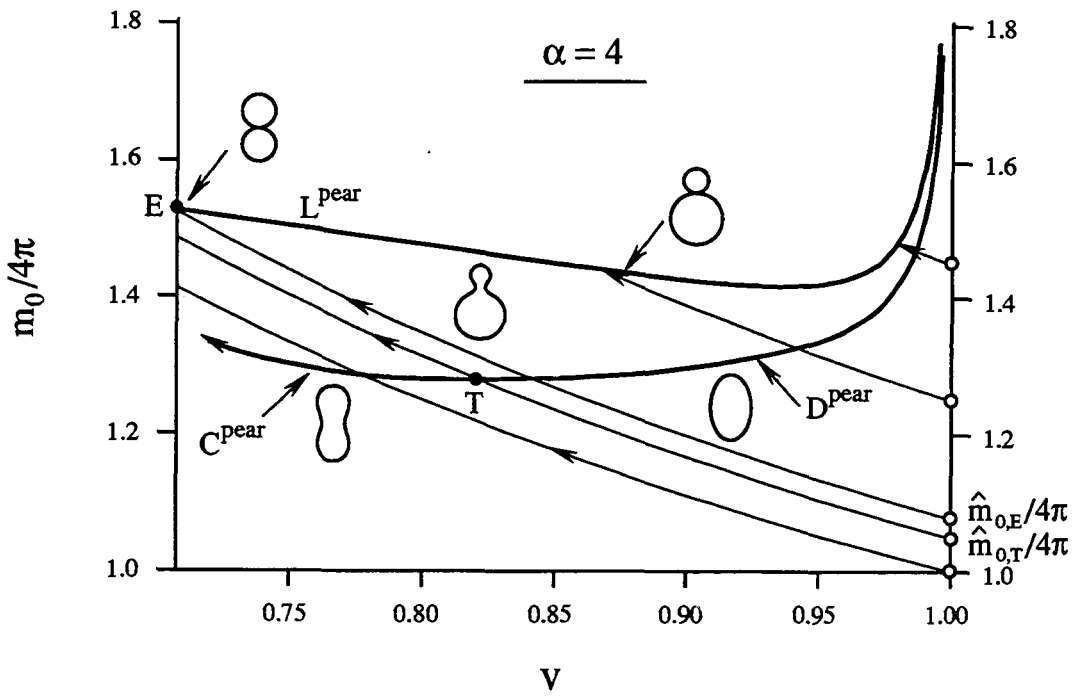


Figure 4.6: Temperature trajectories in the (v, m_0)-plane for $\alpha = 4$. All trajectories are represented by thin lines, with arrows pointing out the direction of a temperature increase starting from a spherical shape but different equilibrium area differences $\hat{m}_0$. For large $\hat{m}_0$, one obtains discontinuous budding with vesiculation; for a small intermediate range of $\hat{m}_0$, discontinuous budding without vesiculation; and, for even smaller $\hat{m}_0$, a continuous budding with reentrant behavior. The larger $\hat{m}_0$, the smaller the vesiculated bud at the limit line L^{pear}.

increasing temperature, however, the vesicle will become either symmetric again (*discontinuous reentrant trajectory*) and/or approach another limiting shape for $v < \sqrt{2}/2$. Since we have not attempted analyzing the shapes for $v < \sqrt{2}/2$, we cannot yet decide between these two alternatives.

(iii) *Continuous budding without vesiculation.* For $\hat{m}_0 < \hat{m}_{0,T}$, the vesicle exhibits a continuous symmetry breaking as it passes into the pear region. Beyond this, either it will become symmetric again (*continuous reentrant trajectory*) and/or it will approach another limiting shape for $v < \sqrt{2}/2$. We have not investigated the lowest value $\hat{m}_0$ for which this trajectory

occurs. For even lower values of $\hat{m}_0$, the trajectories either remain in the prolate/dumbbell phase or reach non-axisymmetric or oblate and stomatocyte shapes.

This picture holds for intermediate values of α, with $\alpha_1 < \alpha < \alpha_2$. The lower bound $\alpha_1 \simeq 1.2$ of this regime is determined by $v_T(\alpha_1) = 0.72$. The reason for making this choice has been given in Section 4.2.1. Unfortunately, a numerical determination of the upper limit α_2 of this regime, defined by $\hat{m}_{0,\mathrm{T}}(\alpha_2) = \hat{m}_{0,\mathrm{E}}(\alpha_2)$, is quite difficult. We were able numerically to follow the tricritical point up to $\alpha = 6.5$, which is still in the intermediate regime. For large α, the asymptotic analysis of Appendix E shows that $\hat{m}_{0,\mathrm{E}} < \hat{m}_{0,\mathrm{T}}$. Thus, we can conclude $6.5 < \alpha_2 < \infty$, but we do not know how tight this lower bound is. In Fig. 4.7, we show how the values $\hat{m}_{0,\mathrm{T}}$ and $\hat{m}_{0,\mathrm{E}}$ and, consequently, the qualitatively different budding and vesiculation trajectories depend on α.

(B) For large values of α ($\alpha > \alpha_2$) both T and E trajectories persist, but with a reverse in the ordering, i.e., $\hat{m}_{0,\mathrm{T}} > \hat{m}_{0,\mathrm{E}}$, as illustrated in Fig. 4.7. Similarly to case (A), there are three scenarios for a thermally-driven shape evolution.

(i) *Discontinuous budding with vesiculation* still occurs, for $\hat{m}_0 > \hat{m}_{0,\mathrm{T}}$. The discussion on the metastability/hysteresis parallels that given for Scenario (i) in case (A).

(ii) *Continuous budding with vesiculation* becomes possible for $\hat{m}_{0,\mathrm{E}} < \hat{m}_0 < \hat{m}_{0,\mathrm{T}}$, as a result of the change of ordering of the T and E trajectories. However, the window in $\hat{m}_0$ for this to happen, i.e., ($\hat{m}_{0,\mathrm{E}}$, $\hat{m}_{0,\mathrm{T}}$), is very narrow (see Fig. 4.7). This scenario will be difficult to observe experimentally.

(iii) *Continuous budding without vesiculation* remains a possibility for $\hat{m}_0 < \hat{m}_{0,\mathrm{E}}$. Upon heating the vesicle will follow a continuous reentrant sequence, in which the vesicle experiences the up/down symmetry breaking first and becomes pear-like, then restores this symmetry and may eventually reach another limiting shape. There exists a smallest value of $\hat{m}_0$, below which the particular budding associated with prolate ellipses and pears is no longer present. However, we have not obtained this value as a function of α systematically (as shown by the shaded region in Fig. 4.7).

(C) For small values of α ($\alpha < \alpha_1$) the situation is less clear, since the evolution of the tri-

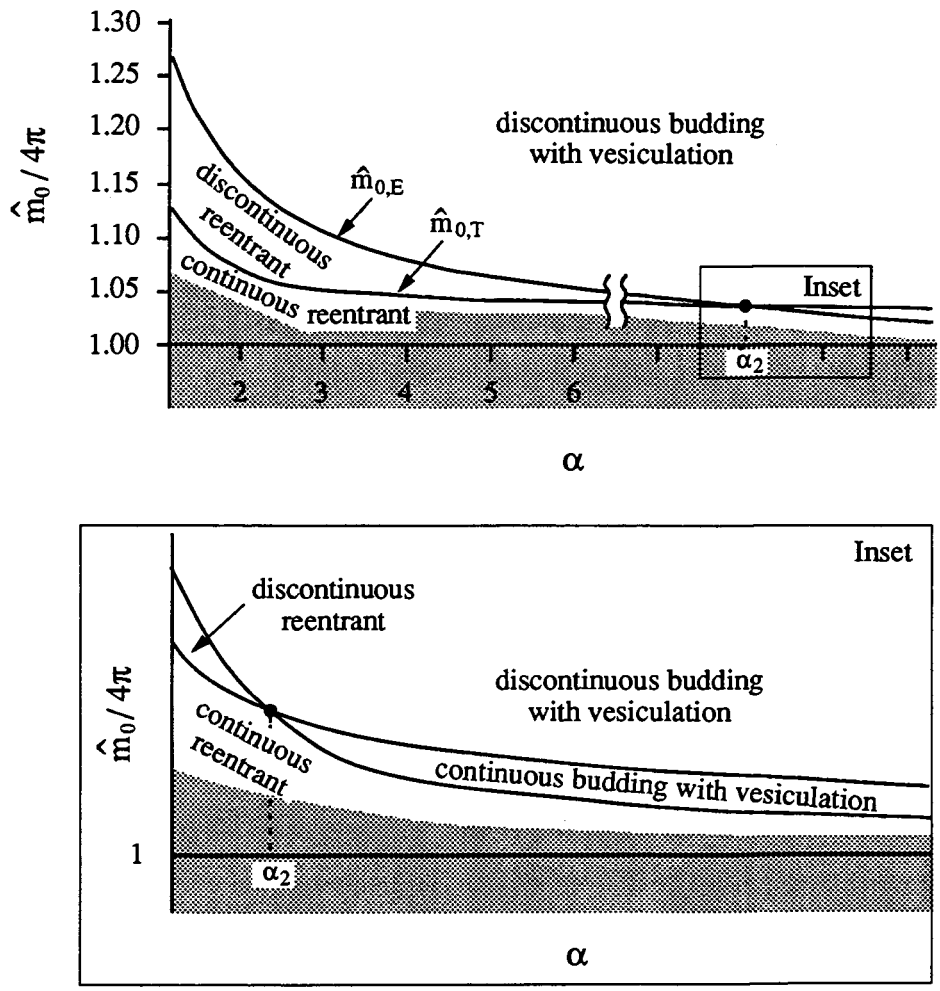


Figure 4.7: Character of the budding as a function of the equilibrium area difference $\hat{m}_0$ and α. For $\alpha > 6.5$ the curve $\hat{m}_{0,T}$ can no longer be obtained numerically. The schematic representation for large α (inset) shows the small region where continuous budding transitions with subsequent vesiculation exist. The gray area shows schematically the region in which the transition from prolates to pears are no longer present.

critical point with α has not yet been investigated[13]. However, for $\hat{m}_0 > \hat{m}_{0,E}$, it is certain that there is discontinuous budding followed by vesiculation. Since the transition to the pear region is first order, we expect hysteresis. For values of $\hat{m}_0$ that are smaller than $\hat{m}_{0,E}$, but immediately follow $\hat{m}_{0,E}$, we expect a reentrant trajectory, in which the transition to the pear region is discontinuous.

In summary, the ADE model exhibits different temperature trajectories even for a fixed value of α depending on the equilibrium area difference $\hat{m}_0$. In particular, continuous budding with subsequent vesiculation occurs only in a restricted range of α and, even there, over a very narrow range of $\hat{m}_0$. These results were obtained under the assumption of zero spontaneous curvature. For non-zero spontaneous curvature, one would have to know the temperature dependence of this quantity, for which there are so far no measurements available. Assuming C_0 to be temperature independent, one obtains still the same scenarios for budding and vesiculation as discussed above, since $C_0 \neq 0$ basically shifts the scale on the m_0-axis, as discussed in Section 4.3. In particular, the presence of a non-zero spontaneous curvature does not alter the value of α_1 below which all budding transitions are discontinuous.

4.4.2 Comparison with the experiments

We close this Section with a few comments on the relation between the predictions of the ADE model and the available experimental observations. It is convenient to structure these remarks in the form of questions and answers:

(a) Can the budding and vesiculation scenarios described in Section 1.2 be explained by the ADE model?

Experiments apparently see both continuous and discontinuous budding transitions, followed by vesiculation. Neither the SC nor the ΔA models could explain this diversity, since each of them contains one type of transition but not the other. As we have seen, the ADE

[13] At this point, the division between the regime A and regime C, $\alpha = \alpha_1$, is somewhat an artificial one, since our analysis is not sufficient to show that there is a clear distinction between the behaviors seen in these two regimes.

model does allow both continuous and discontinuous budding, depending on the parameter α (and with some lesser significance C_0). It even allows continuous and discontinuous transitions for a given single value of α, if the initial area difference m_0 is adjusted appropriately. In this sense, the ADE model is at least qualitatively successful.

On the other hand, closer examination reveals some difficulties. To make a quantitative comparison with the experiments we need to know what value of $\alpha = \bar{\kappa}/\kappa$ to use. In Chapter 2, we estimated this ratio as of order unity for common phospholipids ($\alpha \approx 1.1$ for DMPC and $\alpha \approx 1.4$ for SOPC). The only experimental determination of α that we are aware of was for SOPC+POPC mixed-lipid bilayers and used a tether-pulling technique [39]. This measurement yielded $\alpha \approx 1.1 + 1.6/-0.8$, consistent with our expectation, but with a large range of uncertainty. These values of α are substantially smaller than α_2, so the theory predicts that any temperature-induced budding should be discontinuous. Note that this prediction is independent of both the initial area difference m_0 (see Fig. 4.4) and the value of the spontaneous curvature. Such discontinuous budding has, indeed, been observed [9], as in sequence C of Section 1.2; however, scenario B (Section 1.2) of continuous symmetry breaking followed by a discontinuous shrinkage of the pear neck [20], cannot be explained by the ADE model. Not only is the continuous symmetry-breaking transition not available for such a low value of α but, in addition, a discontinuous transition from weak pears to strong pears or to a vesiculated phase is not available in the ADE model for any value of α.

(b) Is there a possible explanation for the inconsistency between the predictions of the ADE model with $\alpha \approx 1$ and scenario B?

It is not clear at this writing whether this disagreement for budding scenario B is serious or whether it is an experimental artifact. One possibility provides a fairly straightforward explanation of the continuous symmetry breaking: It is not always easy to distinguish between unilamellar and multilamellar vesicles in the laboratory. The reason for this difficulty is that observations are done by light microscopy and, even with sophisticated techniques, it is impossible to resolve distances less than a few tenths of a micron. Thus, all determi-

nations of membrane thickness are inferential and depend on looking for regions of lamellar separation, differences between the behavior of different vesicles from the same preparation, etc. If multilamellar vesicles are present, then these vesicles would have a larger effective value of α: For a bilamellar vesicle, for example, treating both bilayers as elastic fluid sheets (as in Chapter 2) leads to an effective $\alpha_{\rm eff} = \alpha(D_{\rm bilayer}/D)^2$, where α is the single-bilayer value and $D_{\rm bilayer}$ is the interbilayer distance. Suppose $D_{\rm bilayer} = 2D + D_{\rm fluid}$ where $D_{\rm fluid}$ is the thickness of any water layer separating the two bilayers. If $D_{\rm fluid} = 0$, the enhancement of α is a factor of four. For $D_{\rm fluid} > 0$, it could be appreciably larger. For $\alpha > \alpha_2$, there is a range of initial area difference for which continuous budding followed by smooth vesiculation can occur. But, the discontinuous transition observed in scenario B between a weak pear and a narrow-necked or full vesiculated shape still remains unexplained, without some additional mechanism.

Another possible explanation is that the shapes identified in scenario B as stable weak pears were, in fact, slow fluctuations and/or unstable dynamical intermediates. Experimentally, the issue centers on how to distinguish between continuous and discontinuous shape transitions. In practice, the experiments are done [9, 20] by slowly ramping up[14] the temperature, while watching the shape evolution via video microscopy. In the early days temperature ramping rates of 0.3 $^\circ C/minute$ were commonly used. More recently, DMPC experiments have been done with rates of 0.1 $^\circ C/15minutes$ [20]. Ordinarily, shape equilibration after a temperature change takes only a few seconds and fluctuations are weak. Near the transition, however, strong shape fluctuations have been observed, sometimes lasting for periods of several minutes or more [24], so there are clearly some long timescales in the system. Furthermore, the observed shape must, of course, evolve smoothly, as it passes through a transition, whether the transition be first-order or second. Thus, even for discontinuous transitions, weak pears may be observed as dynamical intermediates.

These issues have not yet been adequately explored; however, the crucial questions are: What is responsible for the long timescales and can a changing equilibrium shape be

[14]Once the neck closes, there are strong hysteresis effects (probably connected with short-range van der Waals forces in the region of the (microscopic) neck [45, 59]). Thus, reproducible cycling through the transition has not yet been achieved.

distinguished from a slow dynamical transient? One interpretation of the observations is that the transitions are continuous and that the long timescale is the usual critical slowing down, characteristic of a second-order transition. Another interpretation, which we shall now explore, is that the long timescale is associated with approach to the spinodal line (M^{pro}) and that the observed weak pears are fluctuations and/or dynamical transients. From this perspective, the apparent extended region of weak pears is, in fact, a region of metastable ellipses in which there are long-lived thermally-induced pear fluctuations. In this interpretation, the observed discontinuous transition to a narrow-necked or fully vesiculated state is just the termination of metastability marginally before the spinodal M^{pro}, as a thermal fluctuation finally drives the system over the (small) remaining free-energy barrier and it falls to a low-energy stable equilibrium.

There is some experimental evidence to support this interpretation. In the experiments [60] on DMPC, unexplained fluctuations of the "stable" pears back to the prolate ellipse were occasionally observed. Furthermore, in recent experiments on SOPC [24] long-lived pear fluctuations, lasting as long as 10 minutes, are regularly observed.[15]

On the theoretical side, to explore this alternative we must discuss the characteristic timescales of the system. If the transition is first-order, i.e., hysteretic, there are at least two time scales to be recognized: The symmetric shape, whether it is stable globally or just a metastable state, is subject to asymmetric (pear-like) fluctuations. These fluctuations are normally overdamped by the viscosity of the surrounding fluid, and they have a characteristic relaxation time which we shall call τ_1. Another scale, which we shall call τ_2, is the characteristic time for giant fluctuations to occur, which take the system over the kinetic barrier to a new minimum. When the height of the barrier is large compared to the thermal energy, these two timescales are distinct, with τ_1 the short timescale and τ_2 the long one. A very crude estimate of the typical value of τ_1 is obtained simply by dimensional analysis: $\tau_1 \equiv \eta R_A^3/\kappa \simeq 10 seconds$ [52], where $\eta \simeq 0.01 g/cmsec$ is the viscosity of water, $R_A \simeq 10\mu m$ is the size of the vesicle, and $\kappa = 10^{-12} erg$ is a typical bending rigidity. This is entirely

[15]Interestingly, although these fluctuations go in both "up" and "down" directions, preliminary analysis of the fluctuations indicates that they usually do not quite average to zero, so some source of up/down asymmetry apparently remains [24].

consistent with the typical observed timescale for achieving mechanical equilibrium far from any shape transition. However, close to the spinodal point (and also, of course, close to any real second-order transition), the relaxation time of the asymmetric mode that will initiate the onset of the instability can be much longer than seconds, because the corresponding τ_1 acquires an additional factor $1/\lambda$, where λ is the smallest eigenvalue of the linear stability matrix for small shape fluctuations around equilibrium [61]. This eigenvalue goes to zero at the instability , so τ_1 diverges at the spinodal line and is presumably very small just before M^{pro}, when the metastable barrier is comparable to k_BT. Close to M^{pro}, there is no longer a sharp distinction between τ_1 and τ_2, and, generically, all timescales are comparable to τ_1, at least at the mean-field level. We have made some very preliminary estimates (relying on a Langevin description [62] of the thermal lifetime) of how large this additional factor might be. It is easy to come up with timescales of order 10 min., which would be consistent with the SOPC observations. Certainly a more detailed calculation will be necessary to obtain a reliable estimate.

(c) The ADE model relies on variation of the model parameter $\hat{m}_0$ to explain the diversity of shapes and thermal trajectories seen in experiment. Are the values of this parameter required for explanation of the experiments physically reasonable?

$\hat{m}_0$ is essentially the equilibrium or relaxed initial area difference between the two monolayers of the vesicle. This area difference is set at the time the vesicle is formed. Subsequent to vesicle closure it can only change by lipid flip-flop processes, which we have assumed to be slow normally (see Chapter 1). Initial vesicle closure takes place as water is added to the dry lipid (see Section 1.2), and the mechanism is not understood. Nevertheless, some crude estimates can be made: If the two monolayers (inner and outer) are identical in composition and if the inner and outer fluids are the same, then the area difference is just due to the difference in the number of lipid molecules which each monolayer contains, so $\Delta A_0(T) \equiv (N^{out} - N^{in})a_0(T)$, where $a_0(T)$ is the optimal (relaxed) area per molecule at temperature T, and N^{out} and N^{in} represent the numbers of lipid molecules in the outer and inner monolayers, respectively. For a relaxed spherical vesicle with a radius of $10\mu m$, a

bilayer thickness of $3nm$, and a mean area per molecule of $0.6nm^2$, the average number of lipid molecules in each single monolayer is approximately $N^i \simeq 2 \times 10^9$, whereas the optimal (when both monolayers are unstretched) difference in the number of molecules between the two monolayers is $\Delta N \equiv N^{out} - N^{in} \simeq 10^6$. Of course, initial closure may take place in configurations other than spherical. The relaxed area difference goes generally as CRD, where C is a constant, R is a typical vesicle dimension and D is the bilayer thickness. For the sphere $C = 8\pi$ and $\Delta a_0 = 1$. Other closure geometries lead to values of the constant C which can differ from this by factors of order unity, so we expect Δa_0 of order unity with typical fluctuations of the same order. This is broadly consistent with the range of shapes observed experimentally, as may be seen from Fig. 1.5. Note that ΔN of order 10^6 is much larger than $\sqrt{N}$, so random fluctuations of equal-area monolayers cannot account for the required area difference.

(d) We have already observed (Section 4.1) that precooling a vesicle tends to induce outside budding (when the vesicle is rewarmed). Can this be understood in the context of the ADE model?

These experimental observations [9, 20] relate to vesicles which are cooled so that their surface area shrinks and they become, apparently, tense spheres. Characteristically in this state, all observable shape fluctuation ceases. This is consistent with a state of lateral tension in which the higher elastic energy scale KR^2 (see Section 2.1) comes into play, and the interior pressure rises dramatically. The question, then, is, can we in some way link this increase of interior pressure to an increase in Δa_0, which would, in turn, explain the subsequent exterior budding? An intriguing suggestion of Helfrich [63] provides just such a link. Helfrich proposed in another context that the flow of water through a membrane could produce a parallel transport of lipid along the pressure gradient. We speculate that the interior pressure (due to the lateral stress induced by the cooling) may force the outward permeation of water molecules through the bilayer and may also cause the formation of microscopic pores. According to Helfrich's idea, this water flow could carry lipids from the inner leaf of the bilayer to the outer. Likewise, the existence of pores would also fa-

cilitate lipid flip-flop. Both mechanisms lead to a redistribution of lipid molecules between the leaves of the bilayer in such a way as to increase Δa_0 and, therefore, $\hat{m}_0$, thus favoring outside budding. Using the estimate given in Ref. [63] for the coupling between the water and lipid flows (one lipid molecule is transferred by the transmembrane flow of every 5×10^5 water molecules), we arrive at a rough estimate for this induced lipid flip-flop, which is of the order of 10^6 for a cooling of 10 $^\circ C$ from the unstressed state of the vesicle. As explained in the previous paragraph, this is just the right order of magnitude to produce a significant shape change favoring outside vesiculation of small buds (see Fig. 4.6).

(e) The ADE model clearly has some attractive features but also some drawbacks. What experiments might be crucial to establishing its viability?

First, it should be pointed out that no simultaneous measurements of material parameters and shapes (or shape trajectories) has yet been done. Thus, it would be very useful to be able to measure κ and $\bar{\kappa}$ (e.g., by tether pulling [39]) for some vesicle (thus determining α) and then to observe shapes for the same vesicle as temperature is varied. This would allow a straightforward comparison of observed and predicted shapes, which has not so far been possible. Such a measurement for a variety of vesicles could also put to rest any doubts about whether observed vesicles are unilammellar or multilamellar.

The main outstanding difficulty at this point is the budding sequence B of Section 1.2. First, one should obviously reassess the crucial issue of whether stable pears really do or do not exist in these systems. The focus has to be a unique characterization of these shapes as equilibrium shapes in contrast to slow fluctuations or dynamical transients. A very careful temperature control and an observation time long enough to ensure equilibration of the system at each given temperature will be necessary to resolve this important problem. Experiments which study the dynamics of shape fluctuations in this regime are presently in progress [24].

If α can be measured and if continuous and discontinuous transitions can reliably be distinguished, it will be interesting to see whether the tricritical point can be located directly. In this case, the order of the budding transition would crucially depend on the value of the

parameter $\hat{m}_0$ and sufficient variations of $\hat{m}_0$ for the same vesicle (i.e., at fixed α) would lead to budding transitions of different order. Likewise a systematic investigation of the dependence of the temperature trajectories on the equilibrium area difference $\hat{m}_0$ would lead to more insight, even for the case where the budding transition is always discontinuous. This equilibrium area difference could be controlled either by using the precooling mechanism described above or by forced lipid transfer through a trans-membrane pH-gradient as in Ref. [64]. In such a study one could use the temperature interval necessary for budding and the size of the vesiculated bud as important indicators to locate trajectories on a phase diagram like Fig. 4.4. One could then test the prediction that, the more lipid molecules are transferred to the outer layer, the smaller is the increase in temperature necessary to create budding and the smaller is the radius of the vesiculated bud. We hope that this kind of study may be possible in the not-too-distant future.

Chapter 5

Summary

In this thesis, we have studied theoretically equilibrium shapes and shape transitions in systems of fluid lipid-bilayer vesicles, based on two different models (the spontaneous-curvature model and the area-difference-elasticity model). The objective of the thesis was to provide a theoretical basis for understanding the shapes and shape transitions observed in the controlled experiments, which are briefly reviewed in Chapter 1.

The first model we studied was the spontaneous-curvature model. This model has two material parameters, the bending rigidity κ and the spontaneous curvature C_0. An equilibrium shape in this model is the shape of lowest bending energy under the constraints of fixed area A and volume V. There are three length scales in this model, R_A, R_V (where $A \equiv 4\pi R_A^2$ and $V \equiv 4\pi R_V^3/3$) and $R_0 \equiv 2/C_0$. The parameter regime where $R_A > R_V \sim R_0$ (or $R_A > R_V \sim 1$ in our units) was explored by Deuling and Helfrich [8] and many other groups [26, 40]. Characteristic shapes, such as discocytes and stomatocytes, were found (mostly for negative values of the spontaneous curvature). The recent work of Seifert *et. al.* [27] has revealed the systematics in this parameter regime: For $C_0 \lesssim 0$, the typical low-energy shapes are prolates, oblates, discocytes, stomatocytes, and invaginated shapes. The characteristic shape transitions are discontinuous transitions from prolates to oblates (or discocytes), discontinuous symmetry-breaking (or "inside" budding) and continuous inner vesiculation (or invagination). But, outside budding and vesiculation are absent in this regime. Their study also showed that, when $R_V \sim R_0 > 0$ (the low-volume regime), budding

and vesiculation do occur (consistently with our results). In our study of the spontaneous-curvature model, which is reported in Chapter 3, we have investigated the large-volume regime where $R_A > R_V \gg R_0$ or $R_A > R_V \gg 1$ in our reduced units. Our results can be summarized as follows:
(i) The discontinuous transition from a prolate ellipse to a budded shape is present, always followed by a smooth vesiculation. (ii) "Multiplets", each consisting of two or more segments which touch each other tangentially at the axis and can exchange area and volume in a collective way, can be boundary minima in the space of configurations and, therefore, can compete with the regular Euler shapes (shapes that are solutions to the Euler-Lagrange equations) for being the lowest-energy shape. In particular, for those multiplets which are composed of spherical elements (only two radii, R_A and R_B, for the spheres are allowed), the condition for them to be boundary minima is that $1-(1/R_A+1/R_B) < 0$. (iii) The systematics of equilibrium shapes and shape transitions is summarized in a final phase (shape) diagram (Fig. 3.11). Successive buddings and vesiculations are the dominant transitions in this regime. The physical mechanism underlying the sequential budding is not hard to understand: The Helfrich sphere $R = R_0(= 2)$ costs no bending energy; thus, if a quasisphere with a volume much larger than that of the Helfrich sphere has an excess area of the order of the surface area of the Helfrich sphere, it is energetically favourable for the large vesicle to shed the excess area to form a bud approximately the size of a Helfrich sphere.

Another conceptually similar study of the ΔA model was done by Seifert and co-workers [23, 27], a short review of which is given in Chapter 4. It is easy to see that both the SC and the ΔA models produce same catalog of equilibrium shapes (see Section 4.1). However, the sequence of equilibrium transitions for this model was shown to be different from what is predicted by the SC model. In particular, budding (either "inside" or "outside") is always continuous in the ΔA model in the sense that a budded shape with a narrow neck is approached via a continuous symmetry breaking transition from a prolate ellipse to a pear (wide-necked) shape, followed by smooth shrinkage of the neck.

The controlled experiments performed recently on fluid lipid-bilayer (POPC and DMPC) vesicles (see Chapter 1 for references) provided the test ground for the two distinct types of

shape-transition systematics. It turned out that the predictions of neither of the two models about the shape transitions is fully consistent with the available experimental observations. For the SC model, we see two unsatisfactory facts: First, the diversity of shapes and shape transitions (for example, both the discocyte→ stomatocyte sequence and the prolate ellipse→ pear sequence) observed for samples from the same chemical preparation cannot be explained, since this would require that different samples prepared in exactly the same way have different spontaneous curvature and in the SC model there is no mechanism for a distribution of C_0 values in vesicles that are under exactly same chemical conditions. Secondly, the discontinuous character of budding in this model is not compatible with the observed occurrence of budding via a continuous symmetry-breaking transition (trajectory B in Section 1.2). The ΔA model faces a similar dilemma: It does not provide an explanation for the discontinuous thermal-budding trajectory (trajectory C in Section 1.2) or for the discontinuous shrinkage of the neck in Sequence B. We were then led to believe that the resolution of these discrepancies requires consideration of some additional physical effect or effects.

The area-difference-elasticity (ADE) model (see Chapter 2) became our natural candidate for a generalization of the SC and ΔA models: It contains the additional area-difference elasticity, $\bar{\kappa}$, which allows the monolayer area difference to react elastically. We argued in Chapter 2, based on a microscopic model for the bending moduli, that this new effect is comparable to the bending effect considered in the two extant models (SC and ΔA model). The ADE model is, in fact, an interpolation between the SC and ΔA models, and it has the potential to encompass the different scenarios of shape transitions.

We focused on understanding budding and vesiculation within the context of the ADE model. We found, as we expected, that the ADE model is at least qualitatively successful in the following respect: It does contain both the discontinuous budding and continuous symmetry-breaking transitions, depending on the values of the elastic parameters and on the initial condition of vesicles (i.e., ΔA_0). And, the diversity of the observed shapes is not hard to understand in this model: The parameter ΔA_0 depends on the vesicle closure process and, as we argued in Section 4.4.2, can have a reasonable range of distribution,

leading to two different shape sequences. However, a closer examination reveals problems. The estimated value of the ratio $\alpha \equiv \bar{\kappa}/\kappa$ of the area-difference elasticity is close to 1, a fact supported by experimental measurement of this parameter [39]. The theoretical prediction for vesicles having this α value is that thermally-induced budding should always be discontinuous, followed by a continuous vesiculation. Thus, the same difficulty faced by the SC model exists for this model, as well: This prediction is consistent with the sequence C (Section 1.2), but it does not explain the continuous symmetry-breaking transition in Sequence B.

We have discussed in Section 4.4.2 various possibilities for a consistent explanation of this discrepancy within the ADE model. At present, we favour the following scenario: The observed continuous symmetry-breaking transition and apparently "stable" pear shapes are actually effects of the metastability characteristic of the equilibrium discontinuous transition. The pear shapes are in fact long-lived asymmetric (with respect to up/down symmetry) fluctuations about the (symmetric) metastable prolate ellipse. The sudden narrowing of the neck in a pear shape corresponds to the process in which the system is driven by large thermal fluctuations over the kinetic barrier to the global stable state–the narrow-necked shape or the vesiculated shape. Within this picture the apparently discontinuous vesiculation observed in experiments, as the neck of the pear shrinks suddenly to zero, is really a kinetic process in which the metastable prolate is finally driven over a free-energy barrier and the system falls to a stable vesiculated state.

Further experiments will be required to determine whether the foregoing scenario is, in fact, correct. What is clear now, we feel, is that the previous SC and ΔA models are not adequate to describe the observed diversity of budding and vesiculation processes. The ADE model–with the above interpretation is not inconsistent with the experiments we are aware of.

Appendix A

Derivation of the Euler-Lagrange Shape Equations

As we have discussed briefly in Section 3.2.1, the Euler-Lagrange equations which determine the stationary points of the variational free energy (3.1) can have two different forms, depending on the particular parametrization used to represent shapes geometrically. The two sets of differential equations included in Chapter 3 (the Peterson and Helfrich representations) are equivalent in the sense that they must have exactly the same solution for the same parameter values (σ and p) and boundary conditions. Thus, they can be used interchangeably. Despite this manifest equivalence, Ou-Yang *et al.* have claimed that these two representations of the shape equations are intrinsically different [65]. In this appendix, we shall give derivations of the two representations of the shape equations based on the variational principle. Furthermore, we will clarify the issue raised by Ou-Yang by explicitly showing that one expression maps into the other under the appropriate transformation of variables.

We first derive the Euler-Lagrange equations in the Peterson representation. To perform the variation conveniently, we introduce an arbitrary parameter t in the range [0,1] which simply keeps track of the ordering of the points along the contour[1]. Thus, $t = 0$ denotes

[1]The introduction of the variable t is only to facilitate the variation procedure: The total arclength s of the contour depends on the particular shape and varies when the shape changes. Direct variation of the free energy using the s variable is entirely permissible but requires dealing separately with the endpoint variation $s = 0$ and $s = s_0$. By introducing the parameter t, s may be considered a function of t, and the endpoints are dealt with automatically.

the north pole, while $t = 1$ labels the south pole. The way in which t runs along the curve is entirely arbitrary at this point, a feature sometime referred to a "reparametrization invariance" [66]. The two-dimensional projection of an arbitrary axisymmetric shape is, thus, represented by $[r(t), z(t)]$ (refer to Fig. 3.1). In order that the shape be smooth and properly closed, the following boundary conditions have to be satisfied:

$$r(t)|_{t=0,1} = 0, \quad \left.\frac{dz}{dt}\right|_{t=0,1} = 0\,. \tag{A.1}$$

The arclength $s(t)$ depends on $[r(t), z(t)]$ according to

$$s(t) \equiv \int_0^t \left[\left(\frac{dr}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2\right]^{1/2} \quad \text{or} \quad \frac{ds}{dt} \equiv \left[\left(\frac{dr}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2\right]^{1/2}, \tag{A.2}$$

with the boundary conditions $s(0) = 0$ and $s(1) = s_0$. The angle $\Theta(t)$ is defined in these variables via

$$\cot\Theta(t) \equiv \frac{dz/dt}{dr/dt}, \quad \text{with} \quad \begin{cases} \Theta(t)|_{t=0} = \dfrac{\pi}{2} \\ \\ \Theta(t)|_{t=1} = \dfrac{3\pi}{2}\,, \end{cases} \tag{A.3}$$

$$\text{or, equivalently} \qquad \sin\Theta(t) = \frac{dr/dt}{ds/dt}\,. \tag{A.4}$$

Using the geometrical relationships summarized in Eq. (3.7) , we express the free energy Φ explicitly in terms of the three functions $r(t)$, $s(t)$, $\Theta(t)$, as

$$\Phi/\pi = \int_0^1 dt \left\{ r\left[\frac{d\Theta/dt}{ds/dt} - \frac{\cos\Theta}{r} - C_0\right]^2 + 2\Sigma r + Pr^2\cos\Theta \right\} \frac{ds}{dt}\,. \tag{A.5}$$

Note that the three shape functions $r(t)$, $s(t)$, $\Theta(t)$ are not independent of each other because of the local constraint (A.4). We follow here the approach, employed first by Peterson [42], of incorporating this local constraint into the variational procedure through a "Lagrange-multiplier function" $b(t)$. The complete variational energy then reads,

$$\Psi[\Sigma, P; S, b(t)]/\pi \equiv \int_0^1 dt\ \mathcal{L}\left(r(t), \frac{dr/dt}{ds/dt}, \Theta(t), \frac{d\Theta/dt}{ds/dt}\right)\frac{ds}{dt}\,, \tag{A.6}$$

where

$$\mathcal{L}\left(r(t), \frac{dr/dt}{ds/dt}, \Theta(t), \frac{d\Theta/dt}{ds/dt}\right)$$
$$= r\left[\frac{d\Theta/dt}{ds/dt} - \frac{\cos\Theta}{r} - C_0\right]^2 + 2\Sigma r + Pr^2\cos\Theta + b\left[\frac{dr/dt}{ds/dt} - \sin\Theta(t)\right] \quad . \tag{A.7}$$

Varying Ψ with respect to $r(t)$ subject to the boundary conditions (A.1) leads to the differential equation,

$$\frac{ds}{dt}\frac{\partial\mathcal{L}}{\partial r(t)} - \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial(dr/ds)} = 0 \, . \tag{A.8}$$

The corresponding equation for the variation of $\Theta(t)$ is

$$\frac{ds}{dt}\frac{\partial\mathcal{L}}{\partial\Theta(t)} - \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial(d\Theta/ds)} = 0 \, . \tag{A.9}$$

Variation of Ψ with respect to $s(t)$ turns out to give

$$\delta\Psi|_{\delta s} = \delta s_0\,\mathcal{H}(t=1) - \int_0^1 dt\,\delta s\,\frac{d\mathcal{H}(t)}{dt} = 0 \, , \tag{A.10}$$

where the "Hamiltonian function" $\mathcal{H}(t)$ has the following definition,

$$\mathcal{H}(t) \equiv -\mathcal{L} + \frac{dr}{ds}\frac{\partial\mathcal{L}}{\partial(dr/ds)} + \frac{d\Theta}{ds}\frac{\partial\mathcal{L}}{\partial(d\Theta/ds)} \, . \tag{A.11}$$

The Euler-Lagrange equation for $s(t)$ is $d\mathcal{H}/dt = 0$, because Ψ contains only the derivative of $s(t)$. This implies that $\mathcal{H}(t)$ is a conserved quantity along the shape contour. The fact that the variation of the total arclength, δs_0, is in general non-zero leads to the boundary condition $\mathcal{H}(t=1) = 0$. Since $\mathcal{H}(t)$ is a constant along the curve, this boundary condition indicates that $\mathcal{H}(t)$ is identically zero everywhere on the curve.

Up to this point, the equations are written generally, for an arbitrary parametrization t. It is convenient to choose the particular parametrization $s(t) = s_0 t$. With this choice, $ds/dt = s_0$ and the two differential equations (A.8) and (A.9) reduce to Eqs. (3.11) and (3.12) of the text, while (A.10) becomes the Hamiltonian function, Eq. (3.13).

We now turn to the derivation of the Euler shape equations in the Helfrich representation, where the radial distance r from the symmetry axis is eventually chosen to be the

coordinate parametrizing the shape. We consider a simple case in which a vesicle shape has only one belly point, corresponding to $dz/dr = 0$, and is divided into northern and southern "hemispheres".[2] We then write the total free energy as the sum of the energies for the northern (N) and southern (S) hemispheres. Φ_N has the form

$$\Phi_N/\pi = \int_0^{r_m} \frac{r\,dr}{\sqrt{1-(rC_p)^2}} \left[\left(r\frac{dC_p}{dr} + 2C_p - C_0 \right)^2 - P\,r^2 C_p + 2\Sigma \right] , \tag{A.12}$$

where r_m is the distance from the axis at the belly point. The expression for Φ_S/π is similar. An arbitrary parameter t is now chosen, such that $r(t=0) = 0$ and[3] $r(t=1) = r_m$. We can rewrite Φ_N as

$$\begin{aligned} \Phi_N/\pi &= \int_0^1 dt \frac{r}{\sqrt{1-(rC_p)^2}} \left[\left(r\frac{dC_p}{dr} + 2C_p - C_0 \right)^2 - P\,r^2 C_p + 2\Sigma \right] \frac{dr}{dt} \\ &\equiv \int_0^1 dt\; \mathcal{G}(r(t), C_p(t), \frac{dC_p/dt}{dr/dt})\, \frac{dr}{dt} . \end{aligned} \tag{A.13}$$

The variation of Φ can now be performed by treating the functions $r(t)$ and $C_p(t)$ independently, subject to the boundary condition imposed by the geometry, $r(1) \equiv r_m = 1/C_p(1)$. The variation,

$$\delta\Phi = \delta\Phi|_{\delta C_p} + \delta\Phi|_{\delta r} = 0 , \tag{A.14}$$

leads to the two differential equations,

$$\begin{aligned} &\frac{dr}{dt}\frac{\partial \mathcal{G}}{\partial C_p(t)} - \frac{d}{dt}\left[\frac{\partial \mathcal{G}}{\partial (dC_p/dr)} \right] = 0 \\ &\frac{dr}{dt}\frac{\partial \mathcal{G}}{\partial r(t)} + \frac{d}{dt}\left[\frac{dC_p}{dr}\frac{\partial \mathcal{G}}{\partial (dC_p/dr)} \right] - \frac{d\mathcal{G}}{dt} = 0 \quad , \end{aligned} \tag{A.15}$$

along with the boundary condition at the belly point,

$$\delta C_p(1) \left. \frac{\partial \mathcal{G}}{\partial (dC_p/dr)} \right|_{t=1} - \delta r(1) \left. \left[\frac{dC_p}{dr}\frac{\partial \mathcal{G}}{\partial (dC_p/dr)} - \mathcal{G} \right] \right|_{t=1} = 0 . \tag{A.16}$$

[2]The generalization of the variational procedure demonstrated below to more general shapes with more than one belly point is straightforward.

[3]The reason for using this parameter is, again, that it leads to an efficient treatment of the endpoint variations.

If we choose $r = tr_m$, the differential equations (A.15) reduce to

$$\frac{dC_m}{dr} = \frac{r}{2[1-(rC_p)^2]}\left\{C_p\left[(C_p-C_0)^2 - C_m^2\right] + 2\Sigma C_p - P\right\} - \frac{C_m - C_p}{r} \qquad \text{(A.17)}$$

and the geometrical condition,

$$\frac{dC_p}{dr} = \frac{C_m - C_p}{r}\,. \qquad \text{(A.18)}$$

Eqs. (A.17) and (A.18) constitute the Euler-Lagrange system (3.5) given first by Helfrich [8]. The boundary condition (A.16) reduces to

$$\left\{C_p\left[(C_p-C_0)^2 - C_m^2\right] + 2\Sigma C_p - P\right\}\Big|_{r=r_m} = 0\,. \qquad \text{(A.19)}$$

Finally, we show that, by appropriate transformation of variables, one can pass from one representation of the shape equations to the other: Using the fact that $ds/dr = 1/\sin\Theta$, and $\sqrt{1-(rC_p)^2} = \sin\Theta$ (for the northern hemisphere), we re-express the Helfrich equation (A.17) for C_m in terms of the arclength s,

$$\frac{dC_m}{ds} = \frac{1}{2}\frac{r}{\sin\Theta}\left\{C_p\left[(C_p-C_0)^2 - C_m^2\right] + 2\Sigma C_p - P\right\} - \frac{\sin\Theta}{r}(C_m - C_p)\,. \qquad \text{(A.20)}$$

To reach this result in the Peterson representation, we note that the Lagrange-multiplier function $b(s)$ is fixed via the conservation of the Hamiltonian function to ensure axial symmetry. Explicitly, it takes the form,

$$b(s) = -\frac{1}{\sin\Theta}\left\{r\left[C_m^2 - (C_p-C_0)^2\right] - p\,r^2\cos\Theta - 2\,\Sigma\,r\right\}\,. \qquad \text{(A.21)}$$

Substituting Eq. (A.21) into the differential equation Eq. (3.11) for C_m, we arrive at Eq. (A.20).

Appendix B

Numerical Algorithms for the Shape Calculation

In this Appendix, we describe the algorithms we employed to solve numerically the Euler-Lagrange equations (3.8)–(3.12) for stationary shapes.

At a fixed point in (σ, p) space, there exists a set of solutions (Euler singlets), each of which is characterized by, e.g., its curvature at the north pole[1] $\alpha_0^{(n)}$. The task of finding regular solutions is, therefore, equivalent to locating this set of initial curvatures. Integrating the differential equations (3.8)–(3.12) with some $\alpha_0^{(n)}$ from this set, subject to the additional boundary conditions, $z(0) = 0$, $r(0) = 0$, $\Theta(0) = \pi/2$, and $b(0) = 0$, gives a closed Euler shape.[2] Now, consider a small deviation in initial curvature, $\alpha_0 = \alpha_0^{(n)} + \delta$. We observed that numerical integration then resulted in, typically, one or the other of the two shape contours depicted in Fig. B.1(a). Fig. B.1(b) shows the systematic convergence of integration results towards a closed shape.

Like the north pole, the south pole is a singular point of the equations and cannot be reached numerically. When the typical length scale (R_V) is not too large and numerical convergence of the integration is good, the "shooting method" [67] is used to overcome

[1]The two principal curvatures at the axis for an axisymmetric closed shape are the same.

[2]Since the north pole is a singular point of the shape equations ($r = 0$), the numerical integration cannot in practice be started exactly at the pole. Close to the pole the local solution can be found as series expansions in the arclength s. The integration is thus started at some point s^* (typically 10^{-3}) away from the axis. The boundary values needed for the integration are obtained by evaluating the series expansions at this point.

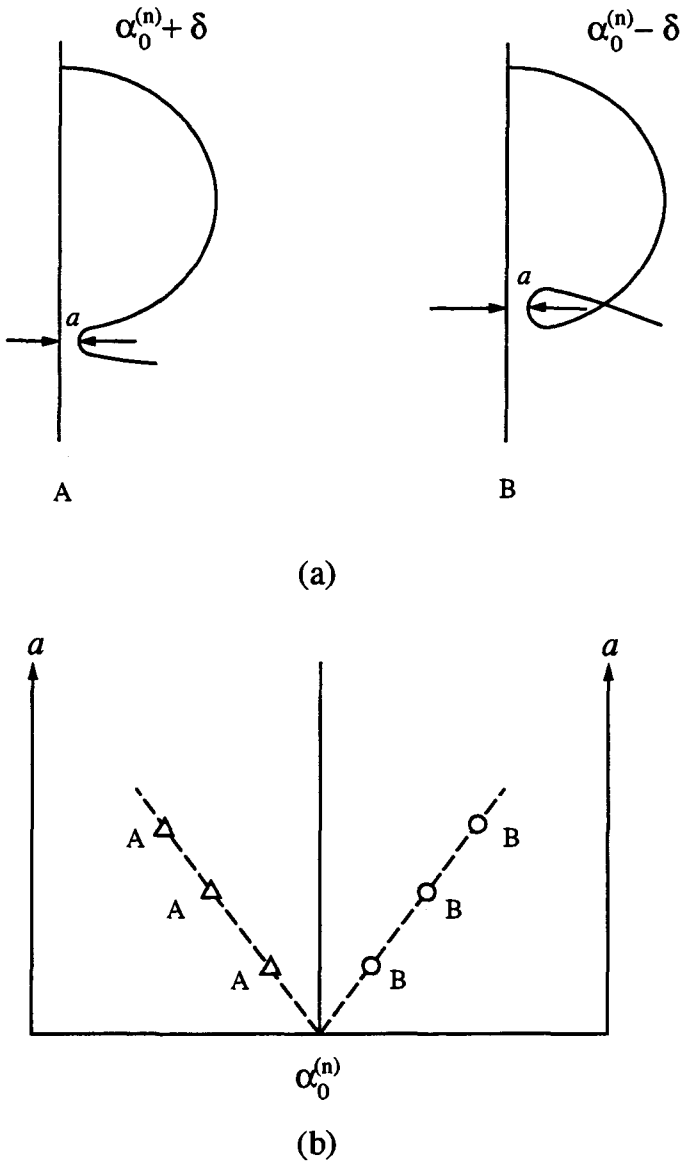


Figure B.1: Schematic illustration of the numerical search for an Euler shape. (a) This shows the two typical behaviors near the axis of a shape, arising from the numerical integration of the Euler shape equations, corresponding to the deviations $\pm\delta$ in the initial curvature from the "perfect value". (b) This shows how successive attempts converge towards an Euler shape, as the deviation in initial curvature gets smaller and smaller.

this difficulty as follows: A reasonable approximation $\tilde{\alpha}_0$ to $\alpha_0^{(n)}$ is picked through the initial search process (see Fig. B.1(b)). Estimates $\tilde{s}_0$ for the total arclength and $\tilde{\alpha}_1$ for the curvature at the south pole are obtained in a similar manner. These estimates serve as trial values in the shooting procedure. In each iteration, we start from the north pole with initial curvature $\tilde{\alpha}_0$ and integrate out to some fixed arclength $\bar{s} < \tilde{s}_0$, thus calculating $r^{(1)}(\bar{s})$, $C_m^{(1)}(\bar{s})$, $\Theta^{(1)}(\bar{s})$, $z^{(1)}(\bar{s})$, and $b^{(1)}(\bar{s})$. Similarly, we perform a second (backwards) integration, starting from the south pole with $\tilde{\alpha}_1$ and integrating out from $\tilde{s}_0$ to $(\tilde{s}_0 - \bar{s})$, thus calculating $r^{(2)}(\bar{s})$, $C_m^{(2)}(\bar{s})$, $\Theta^{(2)}(\bar{s})$, $z^{(2)}(\bar{s})$, and $b^{(2)}(\bar{s})$. For the correct choice of trial parameters ($\tilde{\alpha}_0 = \alpha_0^{(n)}$, $\tilde{\alpha}_1 = \alpha_1^{(n)}$, and $\tilde{s}_0 = s_0^{(n)}$), which characterize an Euler shape, these two sets of calculated values at $\bar{s}$ should match. In practice, $\tilde{\alpha}_0$, $\tilde{\alpha}_1$ and $\tilde{s}_0$ are iteratively adjusted until $r^{(1)}(\bar{s}) = r^{(2)}(\bar{s})$, $C_m^{(1)}(\bar{s}) = C_m^{(2)}(\bar{s})$, and $\Theta^{(1)}(\bar{s}) = \Theta^{(2)}(\bar{s})$. The matching of $z(s)$ and $b(s)$ is automatic, since these quantities depend on r, C_m and Θ.[3]

Once one stationary shape has been successfully found, it is relatively simple to follow its evolution under small changes of σ and p, thus tracing out a particular free-energy sheet over (σ,p) plane.

Finding the appropriate shape for given surface area A and volume V is done by adjusting σ and p. The shooting method described above facilitates this adjustment, since we can supplement the shape equations (3.8)–(3.12) with two more differential equations for A and V,

$$\frac{dA}{ds} = 2\pi r \,, \tag{B.1}$$

and

$$\frac{dV}{ds} = -\pi r^2 \cos\Theta \,, \tag{B.2}$$

subject to the boundary conditions,

$$\begin{aligned} A(0) &= 0, \quad & A(s_0) &= A_1 \,, \\ V(0) &= 0, \quad & V(s_0) &= V_1 \,, \end{aligned} \tag{B.3}$$

[3]It should be noted here that, in looking for a shape with up/down symmetry, the shooting procedure can be simplified: In this case, the matching can be done at the equator. Consequently, there are only two to-be-determined parameters, α_0 and $s_0/2$, subject to the matching conditions at the equator, $\Theta = \pi$ and $b = 0$. This simplification was implemented in practice.

where A_1 and V_1 are the prescribed values for A and V. These two equations then add two more conditions to the matching process, and the free parameters, α_0, α_1, s_0, σ, and p, are all determined together.

Special techniques were needed to deal with those stationary shapes that are so unstable that the shooting method does not converge numerically. One particular example for which the shooting procedure fails is the nearly vesiculated shapes where the segments connected via narrow necks are very different in sizes.[4] We, in these cases, performed a straightforward integration of the differential equations.[5] The search for the stationary shape is done based on the schema shown in Fig. B.1, and the final determination of the "perfect" curvature is obtained by sequentially narrowing down the separation between the initial curvatures corresponding to the two types of behavior illustrated in Fig. B.1(b). The smallest separation we reached is of the order of 10^{-6}. The complete shape, its bending energy, surface area and volume are finally calculated by the direct integration of the augmented set of shape equations to a point very close to the south pole ($r = 10^{-3}$).

One important application of the numerical scheme discussed in the preceding paragraph was in locating numerically the kissing (vesiculation) boundary (see Subsection 3.2.3) for necked shapes consisting of two nearly spherical elements and investigating in detail the shape of the narrow neck. For a fixed value of p, the evolution of an initial necked shape under change of the value of σ was studied. It was found that the neck radius a approaches zero as σ reaches $1/2$. Repeated investigation for several values of p indicated that this was true for all p. This established the kissing boundary numerically. The $\sigma = p$ kissing boundary was numerically located in a similar way. Both of these kissing boundaries were subsequently derived analytically (Appendix C).

The constant volume trajectory for necked shapes near the kissing boundary space was obtained in a way somewhat different from what has been described in the preceding paragraphs: An initial shape was found at some values (σ_1, p_1) with surface area $A(\sigma_1, p_1)$ and volume $V(\sigma_1, p_1)$. The value of p is changed to p_2. Two different values of σ, $\sigma_2^{(a)}$

[4]Interestingly, the larger the size difference is, the more unstable the integration is.

[5]This direct integration may be less efficient than the shooting method, but it is certainly "foolproof."

and $\sigma_2^{(b)}$, are then chosen, such that $V(\sigma_2^{(a)}, p_2) > V(\sigma_1, p_1)$ and $V(\sigma_2^{(b)}, p_2) < V(\sigma_1, p_1)$. Bisection between $\sigma_2^{(a)}$ and $\sigma_2^{(b)}$ finally leads to a σ_2 for which $V(\sigma_2, p_2) = V(\sigma_1, p_1)$. It is in this way that we determined the bending-energy curves $E_N(A, V)$, such as the one shown in Fig. 3.7.

Appendix C

Analytical Treatment of the Nearly Vesiculated Shapes

In Chapter 3, we stated that multiplets (or vesiculated shapes) may, under certain conditions, correspond to (boundary) minima in the space of configurations and compete in energy with stable Euler singlets of the same area and volume. Associated generically with multiplets is another class of stationary shapes–the necked shapes: They have the form of two (or more) compact bodies connected by a narrow but finite neck (or necks), and they satisfy shape equations, e.g., Eq. (3.5), which express the constrained minimization of the total bending energy. These two classes of shapes exist in certain regions of the phase space ((σ,p) or (A,V)). Changing control parameters (σ, p or their conjugates A, V) allows one to pass smoothly between necked and vesiculated shapes. We call this change from the necked to vesiculated shapes a "kissing transition," since it occurs at the boundary where the neck approaches zero and the two bodies joined by the neck touch each other tangentially. Depending upon the interplay between the vesiculated and the necked sheets in the region local to the kissing boundary, a kissing transition may sometimes correspond to an equilibrium phase transition, i.e., "vesiculation," as is seen in the phase diagram Fig. 3.11. In other cases, it only indicates the existence of a bifurcation structure for higher-free-energy sheets.

It is helpful to have in mind the following generic picture of a kissing transition: Suppose, at some given σ and p, that α_0^* is an initial, north-pole curvature such that the shape resulting from integration of the Euler equations closes smoothly at the south pole. It is

possible that there exists another initial curvature $\alpha_0 = \alpha_0^* + \delta$, close to α_0^*, which leads to a properly closed necked shape. δ is then a function of σ and p, and the neck radius a depends on δ. The fact that this necked shape satisfies the Euler equations means that it is at mechanical equilibrium; however, the equilibrium may be either locally stable or locally unstable. A locally stable state may (or may not) correspond to a global minimum; a unstable state cannot. By varying the control parameters σ and p, we shall find that it is possible to drive δ and, in turn, a to zero. In this limit, the necked shape then reaches a well defined limit which is simply a "kissing" multiplet. The condition $\delta(\sigma,p) = a(\sigma,p) = 0$ defines a "kissing" boundary in the (σ,p) plane (or, equivalently, in the (A,V) plane). The free-energy sheets $\Phi_{\text{necked}}(\sigma,p)$ and $\Phi_{\text{vesiculated}}(\sigma,p)$, corresponding to the necked and vesiculated shapes, join along the kissing boundary and are close to one another when $a(\sigma,p)$ is small. When this bifurcation structure involves the lowest-energy sheet, then the kissing transition is a vesiculation transition.

This Appendix is devoted to the study of the region near a kissing boundary. Our purpose is to show generically how small-necked structures can be mechanically stable, low-energy configurations, how nearby vesiculated shapes are boundary minima in configuration space, and how the two types of shapes meet along the kissing boundary. In the process, we will show that $a(\sigma,p) = 0$ is equivalent to the "kissing condition" (Eq. (3.24)),

$$C_A(\sigma,p) + C_B(\sigma,p) = 1 \,. \tag{C.1}$$

where $C_A(\sigma,p)$ and $C_B(\sigma,p)$ are the curvatures of the two connected shapes at the osculating point.

Our treatment has three parts. In Section C.1, we present a treatment based on a variational "toy-model." Despite the fact that it is approximate, this treatment captures in a simple way all of the important physics. In Section C.2, we provide a mathematical derivation of the kissing condition and other results for osculating spheres (i.e., sphere multiplets and spheres joined via narrow necks). In Section C.3, we generalize the results to non-spherical axisymmetric shapes.

C.1 The toy model

We have pointed out in Section 3.3 that, at each point in the allowed region of the (σ, p) plane, there are two stationary spheres with the (exact) equilibrium radii,

$$R_{\pm}(\sigma, p) = \left[(2\,\sigma + 1) \pm \sqrt{(2\,\sigma + 1)^2 - 8p}\right] / 2p\,. \tag{C.2}$$

These two spheres have positive radii when $p > 0$ and $\sigma > -1/2$.

To look at the problem of a narrow-necked shape which is close to two osculating spheres, we adopt a variational trial shape S (see Fig. C.1) in which the two end caps (I and IV) are hemispheres with radii R_1 and R_2, respectively, but the inner regions (II and III) are described by

$$\sin\theta = \frac{1}{R_i + a}\,(r + \frac{R_i\,a}{r}),\ i = 1, 2. \tag{C.3}$$

This shape is chosen to fit the boundary condition $\theta = \pi/2$ both at the equators of the

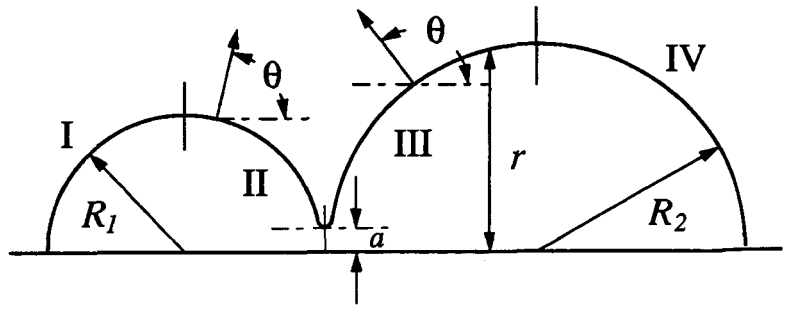


Figure C.1: The "toy shape" used in the variational calculation: Regions I and IV are hemispheres with radii R_1 and R_2; Regions II and III have the form expressed in Eq. (C.3). R_1, R_2 and the neck radius a are the variational parameters of the model.

hemispheres and at the center of the neck. Parametrizing the shape (and in turn the free-energy functional) are three variational variables, R_1, R_2, and a. The free-energy contributions from the spherical regions I and IV are trivial. A brief calculation for region II gives

$$V_{II}(R_1, a) = \frac{\pi R_1^3}{3}\left[(2 + 3\,p_1 + 2\,p_1^2)\,E(q_1) - 3p_1^2\,K(q_1)\right]\,, \tag{C.4}$$

$$A_{II}(R_1, a) = 2\pi\,R_1^2\,(1 + p_1)\,E(q_1)\,, \tag{C.5}$$

and

$$E_{II}(R_1,a) = \frac{1}{2}\left(\frac{2}{R_1+a}-1\right)^2 A_{II}(R_1,a)\,, \tag{C.6}$$

where $p_i = a/R_i$, $q_i = \sqrt{1-p_i^2}$, i=1,2. The functions $E(q)$ and $K(q)$ are complete elliptic integrals [68]. Expanding the elliptic integrals for small a leads to expressions for the volume, area, and free energy in terms of R_1, R_2, and a,

$$V_{I+II}(R_1,a) = \frac{4\pi}{3}R_1^3 + \pi a\, R_1^2 + O(a^2)\,, \tag{C.7}$$

$$A_{I+II}(R_1,a) = 4\pi\, R_1^2 + 2\pi\, a\, R_1 - \pi a^2 \ln(a/R_1) + O(a^2)\,, \tag{C.8}$$

and

$$E_{I+II}(R_1,a) = 2\pi(2-R_1)^2 + \pi a\,[R_1 - 4/R_1] - \pi a^2 \ln(a/R_1)\left[(1-2/R_1)^2/2\right] + O(a^2)\,. \tag{C.9}$$

Equivalent expressions for region III+IV are obtained by replacing R_1 by R_2, so we may finally form the variational free-energy function corresponding to the trial shape of Fig. C.1,

$$\Phi(R_1,R_2,a) = \Phi_0(R_1) + \Phi_0(R_2) + C(\sigma,p;R_1,R_2)\,\pi\, a - D(\sigma,p;R_1,R_2)\pi\, a^2\, \ln a + O(a^2)\,, \tag{C.10}$$

where the coefficients of the terms in a and $a^2\,\ln a$ are

$$C(\sigma,p;R_1,R_2) = (2\sigma+1)(R_1+R_2) - p\,(R_1^2+R_2^2) - 4\,(1/R_1+1/R_2)\,, \tag{C.11}$$

and

$$D(\sigma,p;R_1,R_2) = 2\left[\frac{1}{4}\left(\frac{2}{R_1}-1\right)^2 + \frac{1}{4}\left(\frac{2}{R_2}-1\right)^2 + \sigma\right]\,. \tag{C.12}$$

We now study the behavior of the free-energy sheets at small neck size on the basis of the variational function (C.10). The stationarity conditions $\partial\Phi/\partial R_i = 0$, $i = 1,2$, lead to values for R_1 and R_2 which differ from the solutions $R_\pm(\sigma,p)$ in (C.2) only by terms at most of order a, provided that a is small. Thus, near the stationary points and correct through order a^2, we may replace R_1 and R_2 in Eq. (C.10) by $R_\pm(\sigma,p)$ to write Φ as a function of the neck radius a alone. $\Phi(a)$ plays the role of a Landau function in which a is the order parameter. Evaluating the coefficients $C(\sigma,p;R_1,R_2)$ and $D(\sigma,p;R_1,R_2)$ at

equilibrium sphere radii, $R_{A,B}$, we arrive at an expression for the Landau function,

$$\Phi(a) = \Phi_0(R_A) + \Phi_0(R_B) + 4\pi a \left(1 - \frac{1}{R_A} - \frac{1}{R_B}\right) - D(\sigma, p; R_A, R_B)\, \pi a^2 \ln a + O(a^2)\,, \tag{C.13}$$

where both R_A and R_B may take the value of either of the two sphere radii, $R_\pm(p,\sigma)$. It is clear from Eq. (C.13) that, as long as the coefficient D is positive, the vesiculated shape ($a = 0$) is a boundary minimum of the Landau function whenever $(1/R_A + 1/R_B) < 1$, so it is variationally stable to the formation of a small neck. Conversely, when $(1/R_A + 1/R_B) > 1$, the vesiculated shape becomes locally unstable, and the Landau function has a local minimum at a small but nonzero radius a. This minimum corresponds to a narrow necked shape having the neck size,

$$D(\sigma, p; R_A, R_B)\,(-a \ln a) = 2\left(\frac{1}{R_A} + \frac{1}{R_B} - 1\right)\,, \tag{C.14}$$

with corrections of order a. The kissing boundary is defined by the locus where the bracket on the right hand side of Eq. (C.14) vanishes, i.e.,

$$\frac{1}{R_A} + \frac{1}{R_B} = 1\,, \tag{C.15}$$

which is a special case (i.e., for spheres) of the general kissing condition (C.1). Three cases can be distinguished:

(i) R_A and R_B are different, and correspond to R_+ and R_-, respectively. It follows from (C.2) that the quantity $(1/R_+ + 1/R_-)$ equals unity along the line $\sigma = 1/2$. Eq. (C.14) predicts stable necks in the region just to the right of this line (see Fig. 3.3) with size given by

$$\left[(1+2\sigma)^2 - 4p\right](-a \ln a) = 4(\sigma - 1/2)\,. \tag{C.16}$$

(ii) $R_A = R_B = R_+(\sigma, p) = 2$, which occurs along the line $\sigma = p$ for $\sigma > 1/2$. There are stable small necks in the region just to the left of this line with radii satisfying

$$p(-a \ln a) = \sigma - 1/2 - \sqrt{\frac{(1+2\sigma)^2}{4} - 2p}\,. \tag{C.17}$$

(iii) $R_A = R_B = R_-(\sigma, p) = 2$, which occurs along the line $\sigma = p$ for $\sigma < 1/2$. Stable small necks exist in the region just to the right of this line with radii satisfying

$$p(-a \ln a) = \sigma - 1/2 + \sqrt{\frac{(1+2\sigma)^2}{4} - 2p}\,. \tag{C.18}$$

"Stability" in the preceding paragraph refers, of course, to the free energy Φ in the (σ, p) ensemble. However, it does not necessarily follow that the same stability holds for the mechanical energy E in the (A, V) ensemble, which is simply the Legendre conjugate of the free energy Φ at the stationary point. To study the local energetics there, it is convenient to return to variables A, V. Consider first a vesiculated configuration consisting of two spheres of radii R_1 and R_2 with $R_1 \neq R_2$, so that $V = 4\pi(R_1^3+R_2^3)/3$ and $A = 4\pi(R_1^2+R_2^2)$. Suppose now that a small neck of radius a is allowed to open up between the spheres, so that the overall shape is given by Fig. C.1. In this process the radii R_A and R_B have to deviate from R_1 and R_2 so that A and V stay fixed. Solving Eqs. (C.7) and (C.8) for R_A and R_B as functions of a and using these solutions to calculate the bending energy, we obtain the total mechanical energy as a function of the neck radius,

$$E(A,V;a) - E(A,V;0) = 4\pi \left[1 - (\frac{1}{R_1} + \frac{1}{R_2})\right] a - 2\pi \left[\frac{1}{R_1^2} + \frac{1}{R_2^2}\right] a^2 \ln a + O(a^2) \,. \quad \text{(C.19)}$$

We note that the energy required to make a small neck at fixed A, V vanishes with the neck radius. Clearly, small necks are stable when the condition $(1/R_1 + 1/R_2) > 1$ is satisfied.[1] Mapping back to σ, p variables regenerates Eq. (C.16).

We have now derived, only variationally (for the toy-model shape), the following results: the formula (C.14) for the equilibrium neck radius through order $a \ln a$, the location of the kissing boundary expressed by Eq. (C.15), and the condition $(1/R_1 + 1/R_2) > 1$ for small-neck stability (see Eq. (C.19)). It turns out that Eqs. (C.14) and (C.15) are in fact exact consequences of the Euler shape equations near the vesiculation boundary. We now turn to a more rigorous discussion, which is based on solving the Euler shape equations, in order to put this claim on a firm base.

[1] Any (exact) equilibrium shape has the lowest energy among shapes of the same area and volume. Thus, the minimum energy, (C.19), of toy shapes provides an upper bound to the energy of the necked shapes in any region where these shapes are stable.

C.2 Exact solution for spheres with small necks

For the sake of convenience, we recapitulate here the Euler equations (3.5) expressed in the Helfrich representation,[2]

$$\begin{aligned} \frac{dC_m}{dr} + \frac{dC_p}{dr} &= \frac{r}{2[1-(rC_p)^2]} \left\{ C_p \left[(C_p - 1)^2 - C_m^2 \right] + 2\sigma C_p - p \right\} , \\ \frac{dC_p}{dr} &= \frac{C_m - C_p}{r} . \end{aligned} \tag{C.20}$$

The plan of attack is to develop systematically solutions for the shape in the two regions, far from the neck (where the perturbation theory is valid) and in the (small) neck region (where we employ a scaling Ansatz). Then, by matching those solutions in a region of overlap, we will be able to establish important features of the Euler shapes for $a \to 0$.

Let us focus first on the local shape near the neck in the limit when the neck radius $a \ll 1$. It turns out that a simple scaling Ansatz will allow us to solve the Euler-Lagrange equations (C.20) locally in the neck region order-by-order in terms of the small parameter a. To set this up, we define first the following dimensionless quantities and variables,

$$\begin{aligned} x &\equiv \frac{r}{a} , \qquad \rho \equiv \frac{s}{a} , \\ P(x) &\equiv a\, C_p(r) , \qquad \Psi(x) \equiv a\, C_p(r) + a\, C_m(r) . \end{aligned} \tag{C.21}$$

With this notation, Eq. (C.20) becomes,

$$\begin{aligned} \frac{d\Psi}{dx} &= \frac{x}{2\,(1 - x^2 P^2)} \left\{ P \left[\Psi(2\,P - \Psi) - 2\,a\,P + a^2 \right] + 2\sigma a^2 P - p\,a^3 \right\} , \\ \frac{dP}{dx} &= \frac{\Psi - 2\,P}{x} . \end{aligned} \tag{C.22}$$

[2] Here we use the Helfrich representation of the Euler equations. It is more suitable for analytical treatment than the Peterson representation.

We then propose the simplest scaling Ansatz for a solution to the above shape equations,[3]

$$
\begin{aligned}
P(x) &= P_0(x) + a\,P_1(x) + a^2 P_2(x) + \ldots = \sum_{n=0} a^n P_n(x)\,, \\
\Psi(x) &= a\,\Psi_1(x) + a^2 \Psi_2(x) + \ldots = \sum_{n=1} a^n \Psi_n(x)\,.
\end{aligned}
\tag{C.23}
$$

Substituting (C.23) into the Euler equations (C.22), we generate an infinite sequence of inhomogeneous linear differential equations, corresponding to successive powers of a,

$$
\begin{aligned}
\frac{d\Psi_n}{dx} &= \frac{\Psi_n}{x\,(x^2-1)} + Q_n(\{P_m, \Psi_m; m < n\}; x)\,, \\
\frac{dP_n}{dx} &= -\frac{2\,P_n}{x} + \frac{\Psi_n}{x}\,, \qquad n = 0, 1, 2, 3, \ldots\,.
\end{aligned}
\tag{C.24}
$$

This system can be solved recursively, order by order. The function $Q_n(\{P_m, \Psi_m; m < n\}; x)$ involves solutions from all levels lower than n and is, in general, a cumbersome expression. General solutions to this set of differential equations can be formally found as

$$
\begin{cases}
\Psi_n(x) = K_n \dfrac{(x^2-1)^{1/2}}{x} + \dfrac{(x^2-1)^{1/2}}{x} \displaystyle\int_1^x dy\, \frac{y\,Q_n(y)}{(y^2-1)^{1/2}}\,, \\[2ex]
P_n(x) = \overline{K}_n \dfrac{1}{x^2} + \dfrac{1}{x^2} \displaystyle\int_1^x dy\, y \Psi_n(y)\,,
\end{cases}
\tag{C.25}
$$

where K_n and $\overline{K}_n$ are integration constants. In particular, the solutions of the two lowest orders are easily obtained:

$$
\begin{cases}
\Psi_0(x) = 0\,, \\[2ex]
P_0(x) = \dfrac{1}{x^2}\,,
\end{cases}
\tag{C.26}
$$

[3]This simple scaling Ansatz is certainly consistent not only with the fact that the local mean curvature $(C_m + C_p)$ is not divergent anywhere along the shape contour, but also with the boundary condition $\Psi|_{x=1} = a - \sigma a^2 + O(a^3)$. This boundary condition is the consequence of the condition for the continuity of the shape at $x = 1$, which requires that the curly bracket in (C.22) should be zero at this point.

and

$$\begin{cases} \Psi_1(x) = 1 + K_1 \dfrac{(x^2-1)^{1/2}}{x}\,, \\ \\ P_1(x) = \dfrac{x^2-1}{2\,x^2} + \dfrac{K_1}{x^2}\left[\dfrac{x}{2}\sqrt{x^2-1} - \dfrac{1}{2}\ln(x+\sqrt{x^2-1})\right]\,. \end{cases} \tag{C.27}$$

Note that the boundary conditions $P(x)|_{x=1} = 1$ and $\Psi_1(x)|_{x=1} = 1$ have been satisfied by choosing the integration constants $\overline{K}_0 = 1$ and $\overline{K}_1 = -1/2$.[4] If we take this solution to describe the part of the neck in region II, then the part in region III[5] is then described by the solution,[6]

$$\begin{cases} \tilde{\Psi}_1(x) = 1 - K_1 \dfrac{(x^2-1)^{1/2}}{x}\,, \\ \\ \tilde{P}_1(x) = \dfrac{x^2-1}{2\,x^2} - \dfrac{K_1}{x^2}\left[\dfrac{x}{2}\sqrt{x^2-1} - \dfrac{1}{2}\ln(x+\sqrt{x^2-1})\right]\,. \end{cases} \tag{C.28}$$

The asymptotic behavior of these solutions in the limit $x \to \infty$ (or $r \gg a$) will be important in what follows. We give explicitly the form of $P(x)$ at large x,

$$\begin{aligned} P(x) &= P_0(x) + a\,P_1(x) + O(a^2) \\ &= \frac{1}{x^2} + a\left[\frac{1+K_1}{2} - \frac{2+K_1+2\,K_1\ln 2}{4\,x^2} - \frac{K_1\ln x}{2\,x^2} + O(\frac{1}{x^4})\right] + O(a^2). \end{aligned} \tag{C.29}$$

The functional behavior of these scaling solutions gives the local shape of *any* narrow neck.

We now look away from the neck region and study a perturbative development of the axisymmetric shape equations about closed Euler shapes. This is relevant because a small deviation δ in initial curvature from that of a closed shape leads to a failure of closure at the opposite end, producing a narrow neck. The question we wish to address is: How is the shape modified if we start the integration with a misfit δ in the initial curvature? If the

[4] However, these two boundary conditions are not sufficient to determine K_1. Rather, K_1 can be related to the derivative $d\Psi/d\rho$ at the center in the neck, which parametrizes a family of solutions corresponding to the same neck size a. Other integration constants K_n, $n > 1$, are found to depend upon K_1.

[5] Regions II and III now refer to the neck region of a general narrow-necked shape which has a geometry similar to the toy-model shape in Fig. C.1.

[6] The change of sign in front of K_1 in Eq.(C.28) compared to Eq.(C.27) reflects the fact that the derivative $dx/d\rho = \pm(1-x^2P^2)^{1/2}$ has two different signs for region II and region III, respectively. Only through this change of sign can the derivative dC_m/ds be kept continuous at the center of the neck.

closed Euler shapes are spheres, a family of perturbative solutions parametrized by δ can be derived explicitly by solving the shape equations. We carry through this program below.

An Euler sphere at some fixed σ, p has an equilibrium radius $R(\sigma,p)$ ($R(\sigma,p)$ now denotes either of the two radii, $R_\pm(\sigma,p)$, as given in Eq. (C.2)), so its two principal curvatures are simply $C_m^{(0)}(r) = C_p^{(0)}(r) = 1/R(\sigma,p)$. We now consider a perturbative solution,

$$C_k(r) = C_k^{(0)}(r) + \epsilon_k(r) = 1/R(\sigma,p) + \epsilon_k(r)\ , \quad k = m,p, \tag{C.30}$$

with the boundary condition

$$\epsilon_k(0) = \delta\ . \tag{C.31}$$

Substituting Eq. (C.30) in the Euler equations (C.22), we find the linearized equation for $\epsilon_p(r)$,

$$\left[r\,(1 - r^2/R^2)\,\frac{d^2}{dr^2} + (3 - 2r^2/R^2)\,\frac{d}{dr} - \frac{r}{2R}\,(p\,R^2 - 2)\right]\,\epsilon_p(r) = 0\ . \tag{C.32}$$

A change of variable to $z \equiv r^2/R^2$ reduces this to the hypergeometric equation [69],

$$z\,(1-z)\,\frac{d^2\epsilon_p}{dz^2} + [2 - (A+B+1)\,z]\,\frac{d\epsilon_p}{dz} - AB\,\epsilon_p = 0\ , \tag{C.33}$$

where

$$A + B = 1/2, \qquad \text{and} \qquad AB = \frac{R}{8}\,(p\,R^2 - 2) = \frac{R^2}{8}\,\left[(2\sigma+1) - \frac{4}{R}\right]\ . \tag{C.34}$$

The boundary condition (C.31) picks out the solution which is regular at $z = r = 0$,[7] and we find

$$\epsilon_p(z) = \delta F(A,B,2;z)\ , \tag{C.35}$$

where $F(A,B,2;z)$ is the hypergeometric function. This solution is only valid on the initial (north-pole) side of the belly (region I in Fig. C.1). The function $F(A,B,2;z)$ is finite as $z \to 1$ and can be expanded there in powers of $(1-z)^{1/2}$,

$$\epsilon_p(z) = \frac{\delta\pi^{1/2}}{2}\,\left[\frac{F(A,B,-1/2;1-z)}{\Gamma(2-A)\Gamma(2-B)} + \frac{8}{3}\,(1-z)^{3/2}\,\frac{F(2-A,2-B,5/2;1-z)}{\Gamma(A)\Gamma(B)}\right]\ . \tag{C.36}$$

[7]The hypergeometric equation (C.32) has two linearly independent solutions: One, denoted by $F(A,B,2;z)$, is analytic at $z = 0$ and can be expressed in terms of hypergeometric series for $|z| < 1$. The other is singular at $z = 0$ and is given by $gF(A,B,2;z)\,\ln z + 1/z \sum_{n=0} h_n\,z^n$ where g and h_n are constants.

To continue beyond the belly into region II, an analytic continuation is required, reflecting the fact that the shape in the belly region is entirely smooth. This continuation leads to an expression for the perturbative solution in region II, [8]

$$\epsilon_p(z) = \frac{\delta\pi^{1/2}}{2}\left[\frac{F(A,B,-1/2;1-z)}{\Gamma(2-A)\Gamma(2-B)} - \frac{8}{3}(1-z)^{3/2}\frac{F(2-A,2-B,5/2;1-z)}{\Gamma(A)\Gamma(B)}\right] . \tag{C.37}$$

Having obtained a perturbative solution for the shape in region II, we are now in a position to examine this solution in the neck region. As $(z,r) \to 0$ in region II, the magnitude of the perturbation generally diverges. Indeed, it is precisely this divergence which prevents closure for small $\delta \neq 0$, leading to the formation of a nonzero neck. By developing the solution (C.37) asymptotically for small z (i.e., in the region close to the axis),[9] we arrive at the perturbative solution to the full Euler equations near the neck,[10]

$$\begin{aligned} C_p(r) &= \frac{1}{R} + \delta\left[\frac{\alpha}{z^2} + \frac{1}{2}\beta\ln z + O(1)\right] + O(\delta^2) \\ &= \frac{1}{R} + \delta\left[\frac{\alpha R^2}{r^2} + \beta\ln(\frac{r}{R}) + O(1)\right] + O(\delta^2) , \end{aligned} \tag{C.38}$$

where

$$\alpha = -\frac{2\sin(\pi A)\sin(\pi B)}{\pi(1-A)(1-B)} , \tag{C.39}$$

$$\text{and} \qquad \beta = -\frac{4\sin(\pi A)\sin(\pi B)}{\pi} .$$

[8]The reason for the change of sign is the following: The shape is completely smooth at the belly point. Any shape function (i.e., $r(s)$, $C_m(s)$, and $C_p(s)$) can be expanded in powers of $\Delta s \equiv s - s_b$, $f(\Delta s) = f_b + \sum_{n=1} f_n(\Delta s)^n$. s is the arclength variable, and s_b is the arclength at the belly point. If the shape has the up/down symmetry about the belly point, all the odd-power terms in Δs disappear. If this symmetry is broken, odd-power terms are in general all present. In this case, the shape function $f(\Delta s)$ becomes different on the two sides of the belly point: On one side ($\Delta s > 0$), $f_I(\Delta s) = f_b + \sum_{n=1} f_n|\Delta s|^n$; on the other side ($\Delta s < 0$), $f_{II}(\Delta s) = f_b + \sum_{n=1}(-1)^n f_n|\Delta s|^n$. The coefficients of the odd-power terms in $|\Delta s|$ reverse their sign, in passing from one side of the belly point to the other. This is precisely the sign reversal in Eq. (C.36) and Eq. (C.37), since $(1-z)^{1/2}$ actually corresponds to $|\Delta s|$, and $(1-z)^{3/2}$ is responsible for generating the odd-power terms.

[9]The function $F(A,B,C;1-z)$ becomes divergent as $z \to 0$, if $C-(A+B) = m$, $(m = 0,1,2,...)$.

[10]The perturbative solution $\epsilon_m(r)$ is straightforward to get, using $\epsilon_m(r) = r\,d\epsilon_p/dr + \epsilon_p(r)$.

The combinations $\sin(\pi A)\sin(\pi B)$ and $(1-A)(1-B)$ can be easily evaluated by using Eq. (C.34), and we simply give the results here,

$$\sin(\pi A)\sin(\pi B) = -\frac{1}{2}\cos\left[\frac{\pi}{2}(1+4R-2pR^3)^{1/2}\right],$$

$$\text{and} \qquad (1-A)(1-B) = \frac{1}{8}(pR^3-2R+4). \tag{C.40}$$

The asymptotic shape (C.38) is only valid over a limited range of r, near but not too near the neck. On the one hand, we must have $z \ll 1$ ($r \ll R$), so that higher-order asymptotic terms are negligible. On the other hand, too near the neck the term in $C_p(r)$ linear in δ becomes large and higher-order corrections, δ^2 and larger, would have to be included. To avoid this, we require $\left|\delta\frac{\alpha R^2}{r^2}\right| \ll \frac{1}{R}$. These two requirements together determine an important regime, $\sqrt{|\alpha R^3\delta|} \ll r \ll R$, where Eq. (C.38) well represents the full solution to the Euler equations.

Thus far, we have developed two sets of solutions by looking at two different regions: a family of scaling solutions describing the neck region, which are parametrized by the integration constant K_1 and the neck size a, and a set of perturbative solutions, which depends upon the initial curvature misfits, δ_A and δ_B, at the opposite ends. If there is a narrow-necked shape existing as a solution to the Euler equations, the corresponding K_1, a, and δ_A and δ_B are all functions of σ and p. We expect that matching between these two sets of solutions, if properly done, will uniquely and explicitly determine the parameters. To do this, we form from Eq. (C.38) the combination,

$$\overline{P}(x) \equiv a\,C_p(r) = \frac{a}{R} + \beta\delta a\ln a + \frac{\delta\alpha R^2/a}{x^2} + \beta\delta\, a[\ln x + O(1)] + O(a\delta^2), \tag{C.41}$$

which must be matched to the scaling form (C.29). Matching the constant terms and the $\frac{1}{x^2}$ terms in the equations (C.29) and (C.41) leads to the following two conditions on the parameters,[11]

$$\frac{1}{R_A} + \delta_A\beta_A\ln a + O(a) = \frac{1+K_1}{2}, \tag{C.42}$$

[11] It is not obvious that the use of the asymptotic expression (C.29) in the matching procedure is consistent with the use of the perturbative solution. It turns out, however, that in the end the matching regime, $|\delta|^{1/2} \ll r \ll R$, is also self-consistently the asymptotic regime ($r \gg a$) of the scaling solution, since δ is linear in a for small a.

$$\delta_A \alpha_A R_A^2/a + O(a \ln a) = 1 \,. \tag{C.43}$$

Likewise, from matching between region III and region IV, we have

$$\frac{1}{R_B} + \delta_B \beta_B \ln a + O(a) = \frac{1 - K_1}{2} \,, \tag{C.44}$$

$$\delta_B \alpha_B R_B^2/a + O(a \ln a) = 1 \,. \tag{C.45}$$

Eliminating K_1 from Eqs. (C.42) and (C.44), we obtain again, but now more rigorously, Eq. (C.16) for the equilibrium neck size (to order $a \ln a$). The kissing condition (C.15) is retrieved as the neck radius a approaches zero. From Eqs. (C.43) and (C.45), one sees that the initial misfit for a necked shape is linear in a, but the sign of δ in general depends on σ and p through α and R.

The agreement with the variational calculation seems remarkable, considering how simple the toy shape is. On closer examination, however, this is not at all surprising: The matching presented here involves the $1/x^2$ and constant terms, which are both captured correctly by the toy shape described by Eq. (C.3). The remaining terms in both solutions cannot be easily related at this level of approximation. For example, the $\ln x$ term in the perturbative solution (C.41), has no counterpart in the linear scaling solution. Indeed, the matching of this term can only be done when the nonlinear part of $P(x)$, $a^2 P_2(x)$ is taken into account. We believe that the matching of this term is consistent with the matching of the other two terms.

C.3 Narrow necks between non-spherical Euler shapes

The preceding analysis was performed for the special case where the shapes connected by narrow necks are (nearly) spheres; however, its outcome, namely the equilibrium condition (C.16), and, in turn, the criterion for vesiculation, Eq. (C.15), can be generalized to cases in which the osculating shapes are arbitrary axisymmetric shapes which meet on a common symmetry axis. In the remainder of this Appendix, we shall, develop a basis for this generalization. Broadly speaking, the characteristic asymptotic behaviors of both the

scaling solution and the perturbative solution do not depend on the particular shapes of the osculating objects.

The argument is similar to that given in the preceding Section, only now we consider a perturbation about some arbitrary Euler shape represented by shape functions $\overline{C}_m(r)$, $\overline{C}_p(r)$. Let C_A be the axial curvature of the Euler shape at the osculating point. We write the perturbed shape as

$$C_k(r) \equiv \overline{C}_k(r) + \epsilon_k(r) \,, \quad k = m, p, \tag{C.46}$$

It is then straightforward to show that the linearized differential equation for the perturbation is

$$\frac{d^2\epsilon_p}{dt^2} + \left[\frac{3(1-t^2) - t^2 g_m/2}{t(1-t^2)}\right] \frac{d\epsilon_p}{dt} - \left[\frac{(1-t^2)(g_m + g_p) + 2t^2 f}{2(1-t^2)^2}\right] \epsilon_p = 0 \,, \tag{C.47}$$

where $t \equiv r\, C_A$. The functions g and f are defined as

$$g_p \equiv \left\{ \left[(\overline{C}_p - 1)^2 - \overline{C}_m^2\right] + 2\sigma + 2\,\overline{C}_p\,(\overline{C}_p - 1) \right\} / C_A^2, \qquad g_m \equiv -2\,\overline{C}_p\,\overline{C}_m / C_A^2 \,,$$

$$f \equiv \left\{ \overline{C}_p^2 \left[(\overline{C}_p - 1)^2 - \overline{C}_m^2\right] + 2\sigma \overline{C}_p^2 - p\overline{C}_p \right\} / C_A^4 \,, \tag{C.48}$$

which are all regular. Eq. (C.47) is a second-order linear homogeneous differential equation which has singular points at $t = 0$ and $t = 1$. Close to the axis ($t \to 0$), this equation has two independent solutions [69]: One solution is analytic and can be expressed as a Taylor series in t,

$$\epsilon_p^{(1)}(t) = d_0 \left(1 + A\, t^2 + \ldots\right) , \tag{C.49}$$

where d_0 is the integration constant and $A \equiv \left[(2\sigma + 1) - 4\, C_A\right] / 16\, C_A^2$. The other solution is singular,

$$\begin{aligned} \epsilon_p^{(2)}(t) &= g_0 \left[1 + A\, t^2 + O(t^3)\right] \ln t + \frac{1}{t^2} \left[h_0 + O(t^2)\right] \\ &= \frac{h_0}{t^2} + h_0 \,(1 + 4\, A) \ln t + O(h_0) \,. \end{aligned} \tag{C.50}$$

h_0 is a constant to be determined by some boundary condition. The asymptotic behaviors described by $1/t^2$ and $\ln t$ are universal. This singular solution is generically present in the

neck region and dominates any contribution from the regular solution. We suppose, then, that (C.50) describes the perturbation in a region which is located away from the axis, far enough so that the perturbation can still be considered linear yet close enough so that the asymptotic terms given in Eq. (C.50) pick up the dominant contributions. We expect that further development of this perturbation will lead to a neck as the axis is approached, precisely in the same way as it does when spheres are the osculating objects. The essential feature of the previous matching analysis is again seen here: The explicit perturbative shape to be matched to the scaling asymptotics of the neck has the following form, when scaled by the neck size a:[12]

$$\begin{aligned}
\widetilde{P}(x) &= a\overline{C}_p(r) + a\epsilon_p^{(2)}(r) \\
&= a\left[C_A + c_p^{(2)} r^2 + O(r^3)\right] + \frac{h_0 a/C_A^2}{r^2} + h_0\,(1+4\,A)\,a\,\ln(r\,C_A) + O(a\,d_0) \\
&= \left[a\,C_A + h_0\,(1+4\,A)\,a\ln a + O(a^2)\right] + \frac{h_0/a\,C_A^2}{x^2} + h_0\,(1+4\,A)\,a\ln x + \ldots\,.
\end{aligned} \tag{C.51}$$

Matching the coefficients of the $1/x^2$ and constant terms, as we did in the special case, we arrive at the condition for a stationary neck,

$$B\,(a\,\ln a) = 1 - [C_A(\sigma,p) + C_B(\sigma,p)]\ , \tag{C.52}$$

where the coefficient B is explicitly

$$B \equiv \sigma + \left[C_A(\sigma,p) - \frac{1}{2}\right]^2 + \left[C_B(\sigma,p) - \frac{1}{2}\right]^2\ . \tag{C.53}$$

As the neck size $a \to 0$, the kissing condition (C.1) is again retrieved, now for axisymmetric but non-spherical shapes.[13]

---

[12] $\overline{C}_p(r)$ is analytic close to the axis, corresponding to a smooth closure at the axis. The series expansion for $C_p(r)$ is $\overline{C}_p(r) = C_A + C_p^{(2)} r^2 + O(r^3)$, where $C_p^{(2)}$ is an expansion coefficient that depends on σ, p and the initial curvature C_A.

[13] The function of a global analysis of the perturbed shapes seems to be only in relating the equilibrium neck size to the corresponding misfit(s) in initial curvature.

Appendix D

Calculation of Local Perturbations about Spheres

In this Appendix, we present in detail the calculations which provide the basis for the discussions given in Section 3.3. The focus is on the local bifurcation structures in both (σ, p) and (A, V) spaces, which are associated with the development of perturbations characterized by spherical harmonics of different order $l > 1$.

An arbitrary axisymmetric perturbation about a sphere can be expressed in terms of spherical harmonics, as is in Eq. (3.35),

$$R(\theta) = R\left\{1 + \sum_{l=2}^{\infty} a_l Y_{l,0}(\theta)\right\} \equiv R\{1 + u(\theta)\} \ , \tag{D.1}$$

where the amplitudes $\{a_l\}$ are small. Straightforward calculations lead to expressions for the surface area and the volume of the deformed sphere as functions of R and $\{a_l\}$,

$$\begin{aligned} A(R, \{a_l\}) &= 4\pi R^2 + R^2 \sum_{l=2}^{\infty} a_l^2 \left[1 + \frac{l(l+1)}{2}\right] - \frac{1}{8}\int d\Omega \left(\frac{du}{d\theta}\right)^4 + O(a^5) \ , \\ V(R, \{a_l\}) &= \frac{4\pi}{3} R^3 + R^3 \sum_{l=2}^{\infty} a_l^2 + \frac{R^3}{3} \sum_{l_1, l_2, l_3}^{\infty} G(l_1, l_2, l_3) a_{l_1} a_{l_2} a_{l_3} \ , \end{aligned} \tag{D.2}$$

while the bending energy (2.29) is given by

$$E(R, \{a_l\}) = 2\pi(R-2)^2 + \sum_{l=2}^{\infty} B_l^{(2)}(R) a_l^2 + \sum_{l_1, l_2, l_3}^{\infty} B_{l_1, l_2, l_3}^{(3)}(R) a_{l_1} a_{l_2} a_{l_3}$$

$$+ \sum_{l_1,l_2,l_3,l_4}^{\infty} B^{(4)}_{l_1,l_2,l_3,l_4}(R) a_{l_1} a_{l_2} a_{l_3} a_{l_4} + O(a^5) \,, \tag{D.3}$$

with coefficients

$$B^{(2)}_l(R) = \frac{1}{2}(l-1)l(l+1)(l+2) - Rl(l+1) + \frac{R^2}{2}\left[1 + \frac{l(l+1)}{2}\right] \,, \tag{D.4}$$

and

$$\begin{aligned} B^{(3)}_{l_1,l_2,l_3}(R) \;=\; & \frac{G(l_1,l_2,l_3)}{3}\left\{\frac{(R+2)}{2}[l_1(l_1+1)+l_2(l_2+1)+l_3(l_3+1)]\right. \\ & -\frac{R}{2}[l_1(l_1+1)l_2(l_2+1)+l_2(l_2+1)l_3(l_3+1)+l_3(l_3+1)l_1(l_1+1)] \\ & \left. +\frac{(R-2)}{4}[l_1^2(l_1+1)^2+l_2^2(l_2+1)^2+l_3^2(l_3+1)^2]\right\} \,. \end{aligned} \tag{D.5}$$

The coefficient $G(l_1,l_2,l_3) = \int d\Omega\, Y^0_{l_1}(\Omega)\, Y^0_{l_2}(\Omega)\, Y^0_{l_3}(\Omega)$ is easily expressed in terms of Clebsch-Gordan coefficients and vanishes unless triangle inequalities and parity are satisfied. $B^{(4)}_{l_1,l_2,l_3,l_4}(R)$ is even more cumbersome. For our purposes it is sufficient to give only the term in a_2^4, which is required in extending Eq. (3.40),

$$B^{(4)}_{2,2,2,2}(R) = \frac{15}{56\pi}[-3R^2 + 20R - 24] \,. \tag{D.6}$$

The variational free energy (3.1) therefore has the form given by Eq. (3.36) with the following coefficient for the third-order term:

$$C^{(3)}_{l_1,l_2,l_3}(\sigma,p;R) = B^{(3)}_{l_1,l_2,l_3}(R) - \frac{G(l_1,l_2,l_3)}{3} pR^3 \,. \tag{D.7}$$

The term in a_2^4 is

$$\Phi^{(4)}(\sigma,p;R,a_2) \equiv C^{(4)}_{2,2,2,2}(\sigma,p;R)a_2^4 = \frac{15}{56\pi}[-3(1+2\sigma)R^2 + 20R - 24]a_2^4 \,. \tag{D.8}$$

The spherical-harmonic representation (D.1) of nearly spherical shapes is convenient for the description of bifurcations from the spheres, to which we now turn. For nearly spherical shapes which are mechanically stable, the free energy $\Phi(\sigma,p;R,\{a_l\})$ has stationary points in which all the coefficients $\{a_l\}$ are small (but one or more are nonzero). In the limit $\{a_l \to 0\}$, such shapes can only exist in the neighborhood of the loci defined by $C^{(2)}_l(\sigma,p;R_0(\sigma,p) = 0$,

where $R_0(\sigma, p)$ is one of the radii $R_\pm(\sigma, p)$ of the two stable spheres which exist at each point (σ, p) outside the forbidden region (see Section 3.3 and Fig. 3.3). This condition defines a set of loci in the (σ, p) plane which may be regarded as parametrized $(\sigma_l(R), p_l(R))$ by the radii of the corresponding spheres. For (σ, p) near the locus $(\sigma_l(R), p_l(R))$, there is a nearly spherical stable shape dominated by the perturbation a_l in Eq. (D.1). As (σ, p) moves away from $(\sigma_l(R), p_l(R))$, the asphericity grows and the stable shape develops additional spherical-harmonic content. The free energies of these shapes form sheets over the (σ, p) plane. The sheet labelled by l bifurcates from a sphere sheet along the locus $(\sigma_l(R), p_l(R))$. We now examine the structure of this bifurcation, i.e., the way in which the sheet l separates from the sphere sheet along the bifurcation line. We shall explore this bifurcation structure in detail for the simpler cases, $l = 2$ and $l = 3$.

D.1 Perturbation calculation for $l = 2$

We first consider the $l = 2$ bifurcation from a sphere of radius R_0. Shifts of σ and p from their (spherical) threshold values are denoted by

$$\bar{\sigma}_2 = \sigma_2(R_0) + \epsilon_\sigma, \quad \bar{p}_2 = p_2(R_0) + \epsilon_p \,, \tag{D.9}$$

The difference between R and R_0 is defined as δ_R, i.e.,

$$R \equiv R_0(1 + \delta_R) \,. \tag{D.10}$$

The low-order terms in the free energy relevant for this bifurcation are

$$\begin{aligned}
\Phi(\bar{\sigma}_2, \bar{p}_2; R, \{a_l\}) \;=\; & \Phi_0(\bar{\sigma}_2, \bar{p}_2; R) \\
& + C_2^{(2)}(\bar{\sigma}_2, \bar{p}_2; R) a_2^2 + C_{2,2,2}^{(3)}(\bar{\sigma}_2, \bar{p}_2; R) a_2^3 + C_{2,2,2,2}^{(4)}(\bar{\sigma}_2, \bar{p}_2; R) a_2^4 \\
& + C_4^{(2)}(\bar{\sigma}_2, \bar{p}_2; R) a_4^2 + 3C_{2,2,4,}^{(3)}(\bar{\sigma}_2, \bar{p}_2; R) a_2^2 a_4 + O(a_2 a_4^2) \\
& + C_3^{(2)}(\bar{\sigma}_2, \bar{p}_2; R) a_3^2 + 3C_{2,3,3}^{(3)}(\bar{\sigma}_2, \bar{p}_2; R) a_2 a_3^2 + 3C_{4,3,3}^{(3)}(\bar{\sigma}_2, \bar{p}_2; R) a_4 a_3^2 \\
& + O(a_3^4) \,.
\end{aligned} \tag{D.11}$$

The stationarity condition $\partial\Phi/\partial R = 0$ yields,

$$4\pi\left\{\frac{12}{R_0}\delta_R + (2R_0\epsilon_\sigma - R_0^2\epsilon_p) + (2R_0\epsilon_\sigma - 2R_0^2\epsilon_p)\delta_R - p_2R_0^2\delta_R^2 - R_0^2\epsilon_p\delta_R^2\right\} \tag{D.12}$$

$$+D_2^{(2)}(\bar{\sigma}_2,\bar{p}_2;R)a_2^2 + D_{2,2,2}^{(3)}(\bar{\sigma}_2,\bar{p}_2;R)a_2^3 + \cdots = 0\ ,$$

where the D coefficients denote the derivatives of the C coefficients with respect to R,

$$\begin{aligned} D(\bar{\sigma}_2,\bar{p}_2;R) &\equiv \frac{\partial}{\partial R}C(\bar{\sigma}_2,\bar{p}_2;R) \\ &= D(\sigma_2,p_2;R_0) + O(\delta_R) + O(\epsilon_\sigma) + O(\epsilon_p)\ . \end{aligned} \tag{D.13}$$

The stationarity conditions for the coefficients[1] a_2, a_3 and a_4 give,

$$\frac{\partial\Phi}{\partial a_2} = 2C_2^{(2)}(\bar{\sigma}_2,\bar{p}_2;R)a_2 + 3C_{2,2,2}^{(3)}(\bar{\sigma}_2,\bar{p}_2;R)a_2^2 + 4C_{2,2,2,2}^{(4)}(\bar{\sigma}_2,\bar{p}_2;R)a_2^3 + \ldots = 0\ . \tag{D.14}$$

$$\frac{\partial\Phi}{\partial a_4} = 2C_4^{(2)}(\bar{\sigma}_2,\bar{p}_2;R)a_4 + 3C_{2,2,4}^{(3)}(\bar{\sigma}_2,\bar{p}_2;R)a_2^2 + \ldots = 0\ , \tag{D.15}$$

and

$$\frac{\partial\Phi}{\partial a_3} = 2C_3^{(2)}(\bar{\sigma}_2,\bar{p}_2;R)a_3 + 6C_{2,3,3}^{(3)}(\bar{\sigma}_2,\bar{p}_2;R)a_2a_3 + \ldots = 0\ . \tag{D.16}$$

In order to apply these conditions, we need the coefficients $C^{(n)}$. To the order that we will need them, the required coefficients can be evaluated directly from Eqs. (3.37) and (D.7). These results are:

$$C_2^{(2)}(\bar{\sigma}_2,\bar{p}_2;R) = 0 + (4R_0^2\epsilon_\sigma - R_0^3\epsilon_p) + 4(R_0-3)\delta_R + O(\epsilon\delta_R) + O(\delta_R^2)\ , \tag{D.17}$$

$$\begin{aligned} C_{2,2,2}^{(3)}(\bar{\sigma}_2,\bar{p}_2;R) &\equiv C_{2,2,2}^{(3)}(\sigma_2,p_2;R_0) + \text{small corrections} \\ &= \frac{-4}{21}\sqrt{\frac{5}{\pi}}(5R_0+6) - \frac{1}{21}\sqrt{\frac{5}{\pi}}\left[R_0^3\epsilon_p + 12(2R_0-3)\delta_R\right] + O(\epsilon\delta_R) \end{aligned} \tag{D.18}$$

$$\begin{aligned} C_{2,2,4}^{(3)}(\bar{\sigma}_2,\bar{p}_2;R) &\equiv C_{2,2,4}^{(3)}(\sigma_2,p_2;R_0) + \text{small corrections} \\ &= \frac{-6}{7\sqrt{\pi}}(R_0+32) - \frac{1}{7\sqrt{\pi}}\left[R_0^3\epsilon_p + 2(5R_0-18)\delta_R\right] + O(\epsilon\delta_R)\ , \end{aligned} \tag{D.19}$$

[1] In principle,there is one such equation for every l; however, those modes corresponding to higher l's only come in as high order contributions and do not enter at the order of our treatment.

$$C^{(4)}_{2,2,2,2}(\bar{\sigma}_2,\bar{p}_2;R) \equiv C^{(4)}_{2,2,2,2}(\sigma_2,p_2;R_0) + \text{small corrections}$$
$$= -\frac{15}{14\pi}(2R_0+3) + O(\epsilon) + O(\delta_R)\,, \tag{D.20}$$

$$C^{(2)}_{4}(\bar{\sigma}_2,\bar{p}_2;R) \equiv C^{(2)}_{4}(\sigma_2,p_2;R_0) + \text{small corrections}$$
$$= 126 + O(\epsilon) + O(\delta_R)\,, \tag{D.21}$$

$$C^{(2)}_{3}(\bar{\sigma}_2,\bar{p}_2;R) \equiv C^{(2)}_{3}(\sigma_2,p_2;R_0) + \text{small corrections}$$
$$= 30 + O(\epsilon) + O(\delta_R)\,, \tag{D.22}$$

$$C^{(3)}_{2,2,3}(\bar{\sigma}_2,\bar{p}_2;R) \equiv C^{(3)}_{2,2,3}(\sigma_2,p_2;R_0) + \text{small corrections}$$
$$= -\frac{4}{9}\sqrt{\frac{5}{\pi}}(5R_0+12) + O(\epsilon\delta_R)\,. \tag{D.23}$$

The bifurcation structure can be illustrated most conveniently if we choose a particular trajectory in the (σ,p) plane, $\epsilon_p = 2\epsilon_\sigma/R_0 \equiv \epsilon$, which lies along the straight line corresponding to the sphere of radius R_0. Over this line in the (σ,p) plane the free-energy curve for the sphere is particularly simple,

$$\Phi_S(\bar{\sigma}_2,\bar{p}_2;R) \equiv \Phi_0(\bar{\sigma}_2,\bar{p}_2;R) = \Phi_0(\bar{\sigma}_2,\bar{p}_2;R_0)$$
$$= \Phi_0(\sigma_2,p_2;R_0) + 4\pi R_0^2\epsilon/3 + O(\epsilon^4)\,. \tag{D.24}$$

The stationary solution corresponding to the deformed sphere can be found for this particular choice of σ and p by solving the set of nonlinear equations for R, a_2, a_3, a_4, and so on. From Eq. (D.13), it is easy to see that, to the leading order, δ_R is proportional to a^2. In fact, solving (D.13) for δ_R, we obtain

$$\delta_R = -\frac{R_0}{48\pi}D^{(2)}_2(\sigma_2,p_2;R_0)\,a_2^2 + O(a_2^3) = -\frac{R_0}{48\pi}(4-12/R_0)\,a_2^2 + O(a_2^3) = \frac{3-R_0}{12\pi}\,a_2^2 + O(a_2^3)\,. \tag{D.25}$$

The solution to Eq. (D.15) is found to be

$$a_4 = -\frac{3C^{(3)}_{2,2,4}(\sigma_2,p_2;R_0)}{2C^{(2)}_4(\sigma_2,p_2;R_0)}\,a_2^2 + O(\epsilon^3) = \frac{R_0+32}{98\sqrt{\pi}}\,a_2^2 + O(\epsilon^3)\,, \tag{D.26}$$

which is also quadratic in a_2.[2] The solution to (D.16) in the perturbative regime is $a_3 = 0$. Hence, up to order a_2^3, the variational free energy of this quasispherical shape (resembling an ellipse E) can be evaluated by substituting the above variational solution into (D.11). In this case, we obtain an expression for Φ_E as a function of a_2 alone,

$$\begin{aligned}\Phi_E(a_2) \equiv \Phi_E(\bar{\sigma}_2, \bar{p}_2; R) &= \Phi_0(\sigma_2, p_2; R_0) + 4\pi R_0^2 \epsilon/3 \\ &\quad + 2R_0\epsilon\, a_2^2 - \frac{4}{21}\sqrt{\frac{5}{\pi}}(5R_0+6)\, a_2^3 + O(a_2^4)\,. \end{aligned} \tag{D.27}$$

The function $\Phi_E(a_2)$ plays the role of a Landau function in which a_2 is the order parameter. It is schematically illustrated in Fig. D.1. The stationarity of this Landau function leads to the solution,

$$a_2 = -\frac{4R_0^2\epsilon}{3C^{(3)}_{2,2,2}(\sigma_2, p_2; R_0)} = \sqrt{\frac{\pi}{5}}\frac{7R_0^2\epsilon}{(5R_0+6)} + O(\epsilon^2)\,. \tag{D.28}$$

The free energy at this stationary point thus has the same linear term as that of the sphere (tangent bifurcation), but differs by a term proportional to $(\bar{\sigma}_2 - \sigma_2)^3$, as sketched in Fig. 3.6.

We may now pose a similar question in the (A, V) ensemble: How does the bending energy E of the $l = 2$ prolate ellipse branch behave as function of A and V, when the shape is not too far from a sphere? We choose to consider a constant-volume trajectory, i.e.,

$$V = \frac{4\pi}{3}R_0^3, \quad A = 4\pi R_0^2(1+\Delta)\,. \tag{D.29}$$

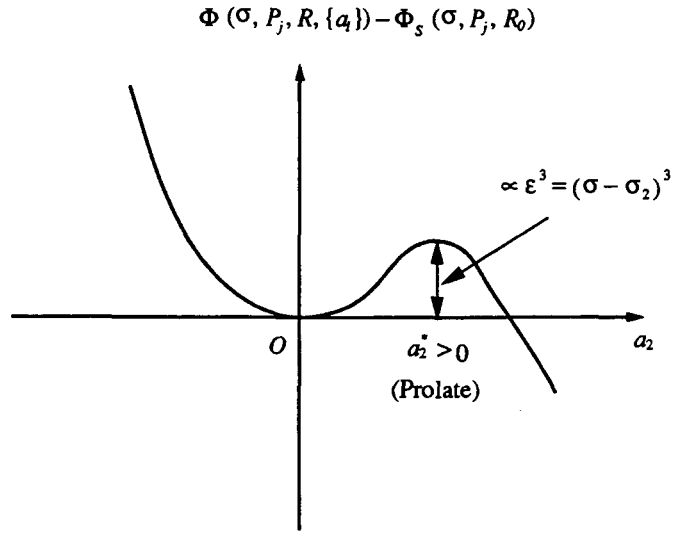
In the perturbative regime, where Δ is small, these two constraints can be rewritten in terms of the deformation variables δ_R and $\{a_l\}$,

$$\begin{aligned}&\delta_R + \delta_R^2 + \delta_R^3 + \frac{1}{4\pi}\left(1 + 3\delta_R + 3\delta_R^2 + \delta_R^3\right)\sum_{l=2}^{\infty} a_l^2 \\ &+\frac{1}{12\pi}\left(1 + 3\delta_R + 3\delta_R^2 + \delta_R^3\right)\sum_{l_1,l_2,l_3}^{\infty} G(l_1,l_2,l_3) a_{l_1}\, a_{l_2}\, a_{l_3} = 0 \quad ,\end{aligned} \tag{D.30}$$

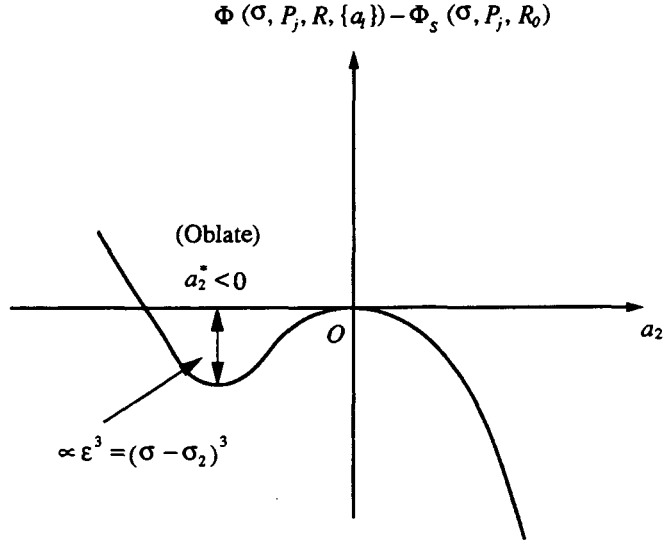
and

$$\begin{aligned}\Delta &= 2\delta_R + \delta_R^2 + \frac{1}{4\pi}\left(1 + 2\delta_R + \delta_R^2\right)\sum_{l=2}^{\infty} a_l^2\left[1 + \frac{l(l+1)}{2}\right] \\ &\quad - \frac{1}{32\pi}\left(1 + 2\delta_R + \delta_R^2\right)\int d\Omega\left(\frac{du}{d\theta}\right)^4 + O(a_l^5)\,.\end{aligned} \tag{D.31}$$

[2] By the same token, $a_6 \propto a_2^3$, since the first term linear in a_6 is $a_2 a_4 a_6$.



(a) $\varepsilon > 0$



(b) $\varepsilon < 0$

Figure D.1: Schematic plot of the Landau function $\Phi_E(a_2)$. For $\epsilon > 0$, $a_2 > 0$ and the shape resembles a prolate ellipse. The corresponding free energy is higher than that of the sphere by a term proportional to $(\bar{\sigma}_2 - \sigma_2)^3$. When $\epsilon < 0$, $a_2 > 0$ and the shape looks like an oblate ellipse, which has a free energy lower than the sphere energy.

From the variational calculation formulated in the (σ, p) ensemble, we have found the quasi-spherical solution described by $R(\sigma, p)$ and $\{a_l(\sigma, p)\}$. In order that the above two constraints be satisfied, ϵ_σ and ϵ_p must be adjusted so that they are appropriate functions of R_0 and Δ. It turns out that, for fixed R_0, $\epsilon_\sigma(R_0, \Delta)$, $\epsilon_p(R_0, \Delta)$, and, thus, $\delta_R(R_0, \Delta)$, $a_l(R_0, \Delta)$ can be expanded in terms of powers of $\Delta^{1/2}$: It is clear that, to lowest order, a_2 is linear in ϵ_σ and ϵ_p (see Eq. (D.14)), while a_4 is quadratic (Eq. (D.15)) and a_6 is at least cubic. Furthermore, ϵ_σ is of the same order of smallness as ϵ_p, and δ_R depends quadratically on ϵ_σ (or ϵ_p): On the one hand, Eq. (D.13) yields the solution $\delta_R = R_0^3(\epsilon_p - 2\,\epsilon_\sigma/R_0)/12 + O(\epsilon_p^2, \epsilon_\sigma^2, \epsilon_p\epsilon_\sigma)$; on the other hand, the volume constraint (D.30) requires that δ_R be quadratic in a_2 and, thereby, in ϵ_σ (ϵ_p). These two conditions together require that the linear term, $(\epsilon_p - 2\,\epsilon_\sigma/R_0)$, in δ_R disappears, i.e., $\epsilon_p = \frac{2}{R_0}\epsilon_\sigma + O(\epsilon_\sigma^2)$. Thus, δ_R becomes quadratic in ϵ_σ (ϵ_p). Using this result and the fact that a_2 is linear in ϵ_σ (ϵ_p), to evaluate the righthand side of Eq. (D.31), we arrive at a polynomial function in ϵ_σ (or ϵ_p) with the leading order term being quadratic. Consequently, the solution to (D.31), $\epsilon_\sigma(R_0, \Delta)$, $\epsilon_p(R_0, \Delta)$ will in general be a polynomial function of $\Delta^{1/2}$.

In what follows we shall carry out the perturbative calculation of the energy $E(V, A) = E(R_0, \Delta)$, through order Δ^2, thus extending Eq. (3.40). The relevant terms are the following:

$$\begin{aligned}
\epsilon_\sigma &= \alpha_\sigma\Delta^{1/2} + \beta_\sigma\Delta + \gamma_\sigma\Delta^{3/2} + \mu_\sigma\Delta^2\,, \\
\epsilon_p &= \alpha_p\Delta^{1/2} + \beta_p\Delta + \gamma_p\Delta^{3/2} + \mu_p\Delta^2\,, \\
\delta_R &= \beta_0\Delta + \gamma_0\Delta^{3/2} + \mu_0\Delta^2\,, \\
a_2 &= \alpha_2\Delta^{1/2} + \beta_2\Delta + \gamma_2\Delta^{3/2} + O(\Delta^2)\,, \\
a_4 &= \beta_4\Delta + O(\Delta^{3/2})\,,
\end{aligned} \tag{D.32}$$

where we have used the fact that $\alpha_0 = 0$ and $\alpha_4 = 0$.

The set of variational equations (D.13)–(D.15) establishes relationships between the

coefficients of σ, p and the coefficients of δ_R and a_l.[3] For example,

$$\begin{aligned}
\alpha_0 &= -(2R_0^2\alpha_\sigma - R_0^3\alpha_p)/12 \\
\alpha_2 &= -\frac{2(4R_0^2\alpha_\sigma - R_0^3\alpha_p)}{3C^{(3)}_{2,2,2}(\sigma_2, p_2; R_0)} \\
\beta_4 &= \frac{R_0 + 32}{98\sqrt{\pi}}\alpha_2^2 \,.
\end{aligned} \tag{D.33}$$

We give the other more complicated relationships only symbolically, to indicate functional dependences,

$$\begin{aligned}
\beta_0 &= \beta_0(\alpha_\sigma, \alpha_p; \beta_\sigma, \beta_p) \\
\beta_2 &= \beta_2(\alpha_\sigma, \alpha_p; \beta_\sigma, \beta_p) \\
\gamma_0 &= \gamma_0(\alpha_\sigma, \alpha_p; \beta_\sigma, \beta_p; \gamma_\sigma, \gamma_p) \\
\mu_0 &= \mu_0(\alpha_\sigma, \alpha_p; \beta_\sigma, \beta_p; \gamma_\sigma, \gamma_p; \mu_\sigma, \mu_p) \,.
\end{aligned} \tag{D.34}$$

The two constraints (D.30) and (D.31) can now be used to determine the coefficients $\{\beta_0, \gamma_0, \mu_0; \alpha_2, \beta_2, \gamma_2; \beta_4(\alpha_2)\}$.[4] We summarize the final results here,

$$\begin{cases}
\beta_0 = -\dfrac{1}{2} \\
\gamma_0 = -\dfrac{\sqrt{10}}{21} \\
\mu_0 = \dfrac{3\,R_0^2 + 120\,R_0 - 5098}{24 \times 7^3} \,,
\end{cases} \tag{D.35}$$

$$\begin{cases}
\alpha_2 = \sqrt{2\pi} \\
\beta_2 = \dfrac{\sqrt{5\pi}}{21} \\
\gamma_2 = -\sqrt{\dfrac{\pi}{2}}\left(\dfrac{81R_0^2 + 4428R_0 - 147440}{36 \times 7^4}\right) \,,
\end{cases} \tag{D.36}$$

[3] To be consistent with the expansion, which is valid to the order of Δ^2, Eq. (D.13) is solved to Δ^2, whereas Eq. (D.14) is solved to $\Delta^{3/2}$, and Eq. (D.15) is solved to Δ.

[4] There are effectively six conditions, corresponding to terms of orders Δ through Δ^2.

and

$$\beta_4 = \frac{\sqrt{\pi}}{49}\,(R_0 + 32)\,. \tag{D.37}$$

The functional forms of ϵ_σ and ϵ_p can be evaluated[5] explicitly to order Δ,

$$\bar{\sigma}_2(R_0;\Delta) = \sigma_2 + \frac{\sqrt{10}(5R_0+6)}{7R_0^2}\Delta^{1/2} + \frac{27R_0^2+76R_0+12612}{1029R_0^2}\Delta + O(\Delta^{3/2})$$

$$\bar{p}_2(R_0;\Delta) = p_2 + \frac{2\sqrt{10}(5R_0+6)}{7R_0^3}\Delta^{1/2} + \frac{2(27R_0^2+1105R_0+6438)}{1029R_0^3}\Delta + O(\Delta^{3/2}) \tag{D.38}$$

Evaluating the bending energy (D.3) at the solution given above, we obtain the additional terms which extend Eq. (3.40) through order Δ^2,

$$-\frac{4\sqrt{10}}{21}(5R_0+6)\Delta^{3/2} - \frac{27R_0^2+76R_0+12612}{1029}\Delta^2 + O(\Delta^{5/2})\,. \tag{D.39}$$

D.2 Perturbation calculation for $l = 3$

The bifurcation associated with the $l = 3$ mode can be treated in a way similar to what we have just done for $l = 2$; however, the absence of an a_3^3 term, as a consequence of the parity of the Clebsch-Gordan coefficients, results in a local bifurcation structure governed by a power law different from the cubic bifurcation we found for $l = 2$. Again, we choose to investigate the local perturbation along the line in the (σ, p) plane corresponding to the sphere of radius R_0, i.e.,

$$\bar{\sigma}_3 = \sigma_3 + \epsilon, \quad \bar{p}_3 = p_3 + \frac{2}{R_0}\epsilon\,. \tag{D.40}$$

The general structure of the free energy Φ is the following:

$$\begin{aligned}\Phi(\bar{\sigma}_3,\bar{p}_3;R,\{a_l\}) \;=\; & \Phi_0(\bar{\sigma}_3,\bar{p}_3;R) \\ & + C_3^{(2)}(\bar{\sigma}_3,\bar{p}_3;R)a_3^2 + C_{3,3,3,3}^{(4)}(\bar{\sigma}_3,\bar{p}_3;R)a_3^4 \\ & + C_2^{(2)}(\bar{\sigma}_3,\bar{p}_3;R)a_2^2 + 3C_{2,3,3}^{(3)}(\bar{\sigma}_3,\bar{p}_3;R)a_2a_3^2 \\ & + C_4^{(2)}(\bar{\sigma}_3,\bar{p}_3;R)a_4^2 + 3C_{4,3,3}^{(3)}(\bar{\sigma}_3,\bar{p}_3;R)a_4a_3^2\end{aligned}$$

[5]There is only one condition relating γ_σ and γ_p and one for μ_σ and μ_p, which is not enough to determine these coefficients.

$$+C_6^{(2)}(\bar{\sigma}_3,\bar{p}_3;R)a_6^2+3C_{6,3,3}^{(3)}(\bar{\sigma}_3,\bar{p}_3;R)a_6a_3^2+\cdots . \quad \text{(D.41)}$$

a_5 and a_8 only contribute to the free energy at higher orders.

The stationarity condition, $\partial\Phi(\bar{\sigma}_3,\bar{p}_3;R,\{a_l\})/\partial R=0$, leads to,

$$\delta_R=-\frac{(5R_0-48)}{48\pi}a_3^2+O(a_3^4)\,. \quad \text{(D.42)}$$

Other stationarity conditions, $\partial\Phi(\bar{\sigma}_3,\bar{p}_3;R,\{a_l\})/\partial a_{2n}=0, n=1,2,3$, determine a_2, a_4, and a_6 as functions of a_3,

$$a_{2n}=-\frac{3C_{2n,3,3}^{(3)}(\bar{\sigma}_3,\bar{p}_3;R_0)}{C_{2n}^{(2)}(\bar{\sigma}_3,\bar{p}_3;R_0)}\,a_3^2+O(a_3^4),\quad n=1,2,3\,. \quad \text{(D.43)}$$

The free energy $\Phi(\bar{\sigma}_3,\bar{p}_3;R,\{a_l\})$ can then be evaluated to give a Landau function of the order parameter a_3,

$$\begin{aligned}\Phi(\bar{\sigma}_3,\bar{p}_3;R,\{a_l\}) &= \Phi_S(\sigma_3,p_3;R_0)+C_3^{(2)}(\bar{\sigma}_3,\bar{p}_3;R_0)\,a_3^2\\ &\quad+\left\{C_{3,3,3,3}^{(4)}(\sigma_3,p_3;R_0)+\Gamma_2+\Gamma_4+\Gamma_6-\frac{(5R_0-48)^2}{48\pi}\right\}a_3^4\\ &\quad+O(a_3^6)\,, \end{aligned} \quad \text{(D.44)}$$

where $C_3^{(2)}(\bar{\sigma}_3,\bar{p}_3;R_0)=5R_0^2\epsilon$ and

$$\Gamma_{2n}=-\frac{9\left[C_{2n,3,3}^{(3)}(\sigma_3,p_3;R_0)\right]^2}{4C_{2n}^{(2)}(\sigma_3,p_3;R_0)},\quad n=1,2,3\,. \quad \text{(D.45)}$$

This Landau function has two symmetrical stationary points, $a_3\propto\pm\sqrt{\epsilon}$, if and only if $\epsilon>0$.[6] The bifurcation is therefore one-sided, and the separation between the sheets goes as a_3^4 or $(\sigma-\sigma_3)^2$.

[6]These two points actually represent the same geometrical shape, reflecting the up/down-symmetry breaking in the shape.

Appendix E

Asymptotic Behavior of the ADE-model Phase Boundaries as $v \to 1$.

In this Appendix, we discuss the asymptotic behavior of the various phase boundaries for the ADE model discussed in Chapter 4 in the limit $v \to 1$.

The analytical calculation of the vesiculation boundary follows from the geometrical constraints associated with the two-sphere limiting shape. Let $r \equiv R_1/R_2$ be the ratio of the radius of the smaller sphere (R_1) to that of the larger sphere (R_2). The reduced volume v defined in Section 4.1 can then be expressed as

$$v \equiv \frac{R_1^3 + R_2^3}{(R_1^2 + R_2^2)^{3/2}} = \frac{1 + r^3}{(1 + r^2)^{3/2}} \,. \tag{E.1}$$

In the limit $v \to 1$, r goes to zero and has the following expansion in powers of $(1 - v)^{1/2}$:

$$r = \sqrt{\frac{2}{3}}(1 - v)^{1/2} + \frac{2}{9}(1 - v) + \sqrt{\frac{1}{6}} \times \frac{65}{54}(1 - v)^{3/2} + O[(1 - v)^2] \,. \tag{E.2}$$

The reduced geometric area difference $m_L(v)$ of this limiting shape is

$$\begin{aligned} m_L(v) &\equiv 4\pi\Delta a = 4\pi \frac{1 + r}{(1 + r^2)^{1/2}} \\ &= 4\pi + 4\pi\sqrt{2/3}(1 - v)^{1/2} + O[(1 - v)] \,. \end{aligned} \tag{E.3}$$

In Chapter 4, we established that $G_n' = 2\,\bar{c}_0$ (Eq. (4.2)), where $\bar{c}_0$ is the dimensionless Lagrange multiplier coupled to the integrated mean curvature M. The Lagrange multiplier

corresponding to the two-sphere limiting shape, $\bar{C}_0$, is found to be $1/R_1 + 1/R_2$, where R_1 and R_2 are the radii of the two kissing spheres, respectively.[1] Making use of this fact, we evaluate G'_a at the vesiculation point and find

$$\begin{aligned} \bar{c}_{0,L}(v) &\equiv \bar{C}_{0,L} R_A = (\frac{1}{R_1} + \frac{1}{R_2})(R_1^2 + R_2^2)^{1/2} \\ &= (1+r^2)^{1/2}(1+\frac{1}{r}) \\ &= \frac{\sqrt{3/2}}{(1-v)^{1/2}} + \frac{2}{3} - \sqrt{\frac{1}{6}} \times \frac{7}{12}(1-v)^{1/2} + O[(1-v)] \,. \end{aligned} \tag{E.4}$$

Eqs. (E.3) and Eq. (E.4) lead straightforwardly to Eqs. (4.27)–(4.29) given in Section 4.2.1.

The phase boundary $\mathrm{D}^{\mathrm{pear}}$ marks the loci of the intersection points of the energy $[W_n(\alpha, v, m_0)]$ curve of the symmetric branch (labeled s) with that of the asymmetric branch (labeled a). Explicitly, the intersection is given by the following condition,

$$\begin{aligned} W_s(\alpha, v, m_{0,D}) &\equiv G_s(v, m_{s,D}) + \frac{\alpha}{2}(m_{0,D} - m_{s,D})^2 \\ = W_a(\alpha, v, m_{0,D}) &\equiv G_a(v, m_{a,D}) + \frac{\alpha}{2}(m_{0,D} - m_{a,D})^2 \,, \end{aligned} \tag{E.5}$$

where $m_{0,\mathrm{D}}$ is related to $m_{s,\mathrm{D}}$ and $m_{a,\mathrm{D}}$ via the mapping (4.13), i.e.,

$$\begin{aligned} m_{0,D} &= m_{s,D} + (1/\alpha) G'_s(v, m_{s,D}) \\ &= m_{a,D} + (1/\alpha) G'_a(v, m_{s,D}) \,. \end{aligned} \tag{E.6}$$

In principle, the loci $m_{0,D}$ of the discontinuous boundary, along with $m_{s,D}$ and $m_{a,D}$, should be exactly solved for from Eq. (E.5) and Eq. (E.6). Typically, $m_{s,D} \neq m_{a,D}$. In the following, we simply give the leading asymptotic behavior of $\mathrm{D}^{\mathrm{pear}}$, found by an approximation: We observe that, as $v \to 1$, $m_{s,D}$ approaches 4π as,[2]

$$m_{s,D} = 4\pi + (4\pi/3)(1-v) + O[(1-v)^{3/2}] \,. \tag{E.7}$$

[1] This is exactly the kissing condition we found in Section 3.2.3. This expression is slightly different from that in Section 3.2.3, which is given in its dimensionless form with the normalization chosen so that $\bar{C}_0 = 1$.

[2] In Appendix E in Ref. [27], it was proved that, for a smooth and deformed sphere, the geometric area difference m has the leading terms given here.

Accordingly, the bending energy G is given by

$$G_s(v, m_s) = 8\pi + O(1-v)\,. \tag{E.8}$$

With these expressions, we write $W_s(\alpha, v, m_{0,D})$ as,

$$W_s(\alpha, v, m_{0,D}) = 8\pi + O(1-v) + \frac{\alpha}{2}\left\{m_{0,D} - 4\pi\left[1 + \frac{1}{3}(1-v) + O[(1-v)^{3/2}]\right]\right\}\,. \tag{E.9}$$

We approximate the bending energy W for the asymmetric branch near the discontinuous transition point by that of the two-sphere vesiculated shape, which is

$$W_a(\alpha, v, m_{0,D}) = 2 \times 8\pi + \frac{\alpha}{2}\left\{m_{0,D} - 4\pi[1 + \sqrt{2/3}(1-v)^{1/2} + O(1-v)]\right\}\,. \tag{E.10}$$

Substituting these two bending energies into Eq. (E.5), we arrive at the leading-order asymptotic behavior of the discontinuous transition boundary,

$$m_{0,D}(\alpha, v) = (2/\alpha)\frac{\sqrt{3/2}}{(1-v)^{1/2}} + O(1)\,. \tag{E.11}$$

Concerning the line $\mathrm{M}^{\mathrm{pro}}$, where the symmetric prolate ellipse develops an instability with respect to a mode breaking the mirror-symmetry, a quantitative assessment of the asymptotic behavior is difficult. Our speculation is that, to leading order, $m_{0,\mathrm{M}}$ also diverges in the form of $const./(1-v)^{1/2}$. For the sake of convenience, we construct our argument within the context of the ΔA-model: In the limit $v \to 1$, the symmetric shape is a slightly deformed sphere, for which the deformations from the sphere can be well represented by the spherical harmonics $Y_{l,0}(\theta, \phi)$ with even l. On the other hand, the first perturbative mode which breaks the mirror symmetry can be described by $Y_{l,0}(\theta, \phi)$ with odd l's. For such nearly spherical shapes, the variational free energy $\Phi[S]$ defined for the ΔA-model as [27]

$$\Phi(\Sigma, P, \bar{C}_0; [S]) = G[S] + \Sigma A[S] + PV[S] - 2\kappa\bar{C}_0 M[S], \tag{E.12}$$

can be expressed in terms of the amplitudes $\{a_l\}$ of the spherical harmonics, as in Appendix D. Varying Φ with respect to $\{a_l\}$ leads to algebraic equations in $\{a_l\}$, which are equivalent to the Euler shape equations. Starting from those values of Σ, P, $\bar{C}_0$ which correspond to the $l = 2$ instability of the sphere [27, 28, 53] and varying these parameters

continuously, we get solutions having nonzero amplitudes for the even spherical harmonics and zero amplitudes for the odd spherical harmonics. These solutions are prolate ellipses with surface area A, volume V and the integrated mean curvature M, given as

$$A = A(\Sigma, P, \bar{C}_0), \quad V = V(\Sigma, P, \bar{C}_0), \quad M = M(\Sigma, P, \bar{C}_0) . \tag{E.13}$$

Eq. (E.13) can be inverted to yield

$$\Sigma = \Sigma(A, V, M), \quad P = P(A, V, M), \quad \bar{C}_0 = \bar{C}_0(A, V, M) . \tag{E.14}$$

We observe that, near the onset of the symmetry-breaking instability, the part of Φ that depends on the amplitudes of the odd-l spherical harmonics takes the form,

$$\Phi(\Sigma, P, \bar{C}_0; \{a_l, l = odd\}) = \sum_{l_1, l_2 = odd} f_{l_1, l_2}(\Sigma, P, \bar{C}_0) a_{l_1} a_{l_2} + O(a_{l_1} a_{l_2} a_{l_3}) . \tag{E.15}$$

At the onset of the instability, the smallest eigenvalue of the coefficient matrix $f_{l_1,l_2}(\Sigma, P, \bar{C}_0)$ changes sign from being positive to being negative. At this point a solution for the $\{a_l\}$ with nonzero values for odd l's first develops. The matrix elements $f_{l_1,l_2}(\Sigma, P, \bar{C}_0)$ in Eq. (E.15) contain only the scaling form[3] $(\bar{C}_0 R_A)(1-v)^{1/2}$. It is then reasonable to anticipate that the condition on the smallest eigenvalue at the onset of the instability will lead to an equation of the form,

$$(\bar{C}_0 R_A)(1 - v)^{1/2} = \text{constant} . \tag{E.16}$$

Following from Eq. (4.4), we have

$$(2\bar{C}_0 R_A) = \partial G_s / \partial m . \tag{E.17}$$

The above condition, combined with the mapping (4.13) in Section 4.1, then implies

$$m_{0,\mathrm{M}} \equiv m_{\mathrm{C}} + \frac{1}{\alpha} G_s \sim \frac{const.}{\alpha(1-v)^{1/2}} . \tag{E.18}$$

We also performed a numerical analysis of this instability for reduced volumes up to $v = 0.994$. The numerical data are consistent with Eq. (E.18). Furthermore, the numerical

[3]This fact can be traced back to the perturbative development of the $l = 2$ symmetric ellipse, which can be expressed in powers of the fractional excess area $\Delta^{1/2}$ (see Appendix D). Note that Δ is linear in $(1-v)$ to lowest order, as $v \to 1$.

analysis shows that *const.* $\approx 2(2.8) > 2(\sqrt{3/2})$; so, the spinodal line M^{pro} has the same asymptotic divergence $[1/(1-v)^{1/2}]$ as the discontinuous boundary and the vesiculation boundary, when $v \rightarrow 1$, but diverges faster than those two boundaries. This implies that the spinodal line extends beyond the vesiculation boundary as v approaches 1.

Bibliography

[1] M. Bessis, *Living Blood Cells and Their Ultrastructure* (Springer-Verlag, New York, 1973).

[2] Theodore L. Steck, *Cell Shape: Determinants, Regulation, and Regulatory Role,* edited by D. Stein Wilfred and Felix Bronner (Academic Press, San Diego, 1989).

[3] R. Lipowsky, Nature **349**, 475(1991).

[4] *Red Cell Membrane: Structure and Function,* edited by G. A. Jamieson, and Tibor J. Greenwalt (J. B. Lippincott Company, U.S.A, 1969).

[5] M. Bloom, E. Evans, and O. G. Mouritsen, Quart. Rev. Biophys. **24**, 3 293(1991).

[6] P. B. Canham, J. Theor. Biol. **26**, 61(1970).

[7] W. Helfrich, Z. Naturforsch. **28c**, 693(1973).

[8] H. J. Deuling and W. Helfrich, J. Phys. (Paris) **37**, 1335(1976).

[9] H.-G. Döbereiner, W. Rawicz, M. Wortis, and E. Evans, (unpublished).

[10] L. Peliti and S. Leibler, Phys. Rev. Lett. **54**, 690(1985).

[11] G. Cevc and D. Marsh, *Phospholipid Bilayers: Physical Principles and Models* (Wiley, New York, 1987).

[12] Charles Tanford, *The Hydrophobic Effect: Formation of Micelles and Biological Membrane* (Wiley, U.S.A., 1980).

[13] J. N. Israelachvili, *Intermolecular and Surface Forces* (Academic Press, London, 1985).

[14] Derek Marsh, *Handbook of Lipid Bilayers* (CRC Press, U.S.A, 1990).

[15] G. Lindblom, H. Wennerström, and G. Arvidson, Int. J. Quant. Chem. **12**, Suppl. 2 153(1977).

[16] R. Homan and H. J. Pownall, Biochim. Biophys. Acta **938**, 155(1988).

[17] J. L. Hutson and J. A. Higgins, Biochim. Biophys. Acta **835**, 236(1985).

[18] E. Evans and W. Rawicz, Phys. Rev. Lett. **64**, 2094(1990).

[19] E. Sackmann, H. P. Duwe, K. Zeman, and A. Zilker, *Structure and Dynamics of nucleid Acids, Proteins and Membranes,* edited by E. Clementi and S. Chin (Plenum, New York, 1986), p.251.

[20] J. Käs and E. Sackmann, Biophys. J. **60**, 825(1991).

[21] M. Mutz and D. Bensimon, Phys. Rev. A **43**, 4525(1989).

[22] E. Evans and D. Needham, J. Phys. Chem. **91**, 4219(1987).

[23] K. Berndl, J. Käs, R. Lipowsky, E. Sackmann, and U. Seifert, Europhys. Lett. **13**(7), 65(1990).

[24] Private communication with Hans-Günther Döbereiner.

[25] S. Svetina, A. Ottova-Lietmannová, and R. Glaser, J. Theor. Biol. **94**, 13 (1982); S. Svetina and B. Žekš, Biomed. Biochim. Acta **42**, S86(1983); **44**, 979(1985).

[26] S. Svetina and B. Žekš, Euro. Biophys. J. **17**, 101(1989).

[27] U. Seifert, K. Berndl, and R. Lipowsky, Phys. Rev. A **44**, 1182(1991).

[28] L. Miao, B. Fourcade, M. Rao, M. Wortis and R. K. P. Zia, Phys. Rev. A **43**, 6843(1991).

[29] M. Wortis, U. Seifert, K. Berndl, B. Fourcade, L. Miao, M. Rao, and R. K. P. Zia, Proceedings of the workshop on *Dynamical Phenomena at Interfaces, Surfaces and Membranes,* Les Houches 1991, edited by D. Beysens, N. Boccara, and G. Forgacs (NovaScience).

[30] E. Evans, Biophys. J. **14**, 923(1974).

[31] S. Svetina, M. Brumen, and B. Žekš, Stud. Biophys. **110**, 177(1985).

[32] E. Evans and R. Skalak, *Mechanics and Thermodynamics of Biomembranes* (CRC, Boca Raton, 1980), p.251.

[33] R. Cantor, *Macromolecules* **14**, 1186(1981)

[34] S. Milner and T. A. Witten, J. de Physique **49**, 1951(1988).

[35] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity* (Pergamon, New York, 1970).

[36] S. Leibler, *Statistical Mechanics of Membranes and surfaces,* edited by D. Nelson, T. Piran, and S. Weinberg, (World Scientific, Singapore, 1989), p.45.

[37] W. Helfrich, Liq. Cryst. **5**, 1647(1989).

[38] D. A. Huse and S. Leibler, J. de Physique **49**, 605(1988).

[39] R. Waugh, J. Song, S. Svetina, and B. Žekš, Biophys. J. **61**, 974(1992).

[40] J. T. Jenkins, J. Math. Biol. **4**, 149(1977).

[41] M. A. Peterson, Mol. Cryst. Liq. Cryst. **127**, 257(1985).

[42] M. A. Peterson, J. Math. Phys. **26**, 711 (1985),

[43] C. F. Schmidt, Y. Barenholz, C. Huang, and T. E. Thompson, Nature **271**, 775(1978)

[44] M. P. Sheetz and S. J. Singer, J. Membr. Biol. **69**, 113(1982).

[45] R. Bruinsma, J. Phys. (Paris) Colloq. **51**, C7-53(1990).

[46] Reinhard Lipowsky, "Domain-induced budding of fluid membranes", preprint (1992); Udo Seifert, "Curvature-induced lateral phase segregation in two-component vesicles", preprint (1992).

[47] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson, *Molecular Biology of the Cell* (Garland Publishing Co., New York, 1989).

[48] E. Sackmann, H. P. Duwe, and H. Engelhardt, Faraday Discuss. Chem. Soc. **81**, 281(1986)

[49] J. C. Luke and J. I. Kaplan, Biophys. J. **25**, 107(1979)

[50] H. P. Duwe, J. Käs, and E. Sackmann, J. de Phys. **51**, 945(1990).

[51] M. A. Peterson, J. Appl. Phys. **57**, 1739(1985).

[52] S. T. Milner and S. A. Safran, Phys. Rev. A **36**, 4371(1987).

[53] Ou-Yang Zhong-can and W. Helfrich, Phys. Rev. Lett. **59**, 2486(1987); **60**, 1209(1987) Phys. Rev. A **39**, 5280(1989)

[54] S. A. Safran, J. Chem. Phys. **78**, 2071(1983)

[55] M. J. Sewell, *Maximum and Minimum Principles* (Cambridge University Press, Cambridge, England, 1987), p.98ff

[56] F. A. Nezil and M. Bloom, Biophys. J. **61**, 1176 (1992).

[57] M. R. Morrow, J. P. Whitehead, and D. Liu, (unpublished, to appear in Biophys. J.).

[58] J. F. Nagle and D. A. Wilkinson, Biophys. J. **23**, 159 (1978).

[59] J. Käs, E. Sackmann, S. Svetina and B. Žekš, "Thermally induced budding of phospholipid vesicles—a discontinuous process", preprint (1992).

[60] Private communication with Hans-Günther Döbereiner and Josef Käs.

[61] M. C. Cross and P. C. Hohenberg, *Pattern Formation Outside of Equilibrium* (in press)

[62] C.W. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, 1985).

[63] E. Boroske, M. Elwenspoek, and W. Helfrich, Biophys. J. **34**, 95(1981).

[64] E. Fargé and P. Devaux, Biophys. J. **92**, 347(1992).

[65] Ou-Yang Zhong-Can, "On shape equations of the axisymmetric vesicles", preprint (1992)

[66] Francois David, *Statistical Mechanics of Membranes and Surfaces,* edited by D. Nelson, T. Piran, and S. Weinberg, (World Scientific, Singapore, 1989), p.158.

[67] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling, *Numerical Recipes: The Art of Scientific Computing* (Cambridge, New York, 1986), pp.582–588.

[68] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products* (Academic Press, New York, 1965), See Formulas 8.111–8.114.

[69] Moorse and Feshbach, *Mathematics for Theoretical Physics* (Academic Press, New York, 1965), See Formulas 8.111–8.114.