January 2010

# Equivariant Deformations of Horospherical Surfaces

Michael Deutsch
*Washington University in St. Louis*

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Mathematics

Dissertation Examination Committee:
Gary Jensen, Chair
Quo-Shin Chi
Steve Fazzari
Renato Feres
Zohar Nussinov
Edward Wilson

# Equivariant Deformations of Horospherical Surfaces

by

Michael Benjamin Deutsch

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment for the degree
of Doctor of Philosophy

May 2010

Saint Louis, Missouri

# Preface

This work is broadly concerned with constant mean curvature (CMC) surfaces in the space forms, particularly the two subclasses of minimal surfaces in Euclidean space and CMC1 surfaces in hyperbolic space, which are known to be locally equivalent via an application of Bonnet's Theorem (often called Lawson's correspondence). CMC surfaces are themselves a subclass of the "isothermal surfaces," a natural class of classical interest in both mathematics and physics, the theory of which combines a surprisingly elegant mixture of ideas from Riemannian, conformal, and Möbius geometry as well as complex analysis and PDE theory. We will indeed profit from a number of these ideas as we study the special case of "horospherical surfaces," a recently-coined term which combines our two subclasses of interest. The focus is on the transformation theory of these surfaces as a way of understanding the general correspondence. The basic original results presented here can be summarized as follows: The classical Goursat transform for minimal surfaces can be interpreted as conformal transformation of the Gauss map, thereby allowing us to "bend" these surfaces in way that is well-suited for certain geometric purposes. This deformation has a simple analogue for CMC1 surfaces which, when properly defined, makes the Goursat transform equivariant with respect to the Lawson correspondence, and greatly increases the number of explicitly computable examples of minimal/CMC1 cousin pairs. A "quaternionic upper-half space" model for hyperbolic 3-space is introduced and it is argued that this is the simplest context in which to understand the correspondence itself. Finally, there is a compelling analogy between (a) the Goursat transformation law and integrability conditions for the "spin curve" of a horospherical surface, and (b) the Lorentz transformation law and equations of motion for the spin wavefunction of a massless fermion. The geometry of spin curves is analyzed by moving frames, and on the basis of the above analogy, it is suggested that horospherical surfaces could have applications to particle physics and vice versa.

Following Hardy, it may also be appropriate to add a brief apology. The reader already familiar with classical surface theory must excuse the rather elementary nature of the remarks in the introduction, a tone perhaps ill-suited to the task of dissertation writing, which more often assumes a format closer to that of an research article. This choice of style is entirely deliberate, in an effort to make the results accessible even to a curious graduate student, and is motivated by the following personal experience: Having receiving what by current standards constitutes a reasonable graduate exposure to differential geometry, I was surprised to discover my ignorance of a huge number of basic classical accomplishments, and I've since learned that this experience is not uncommon among my colleagues. A solid introduction to the classical differential geometry of curves and surfaces seems to have been removed from many graduate curricula, presumably in order to engage advanced topics as quickly as possible and help launch students into research directly. While this strategy appears to meet that objective, it may do some disservice to graduate education itself, not only because it leaves a slight deficiency in general knowledge, but also because a good deal of conceptual intuition and continuity of the subject will be lost on the student who lacks a sense of the natural historical progression of its ideas. Additionally, students tend to miss out on what would otherwise be exciting potential research directions, and I count myself lucky to have found an advisor with considerable knowledge and interest in active concerns of the classical theory. The exposition presented here aims to help fill this "classical gap," the source of which (as we perceive and describe it) suggested our rather paradoxical approach: we use rhetoric which assumes familiarity with, say, Riemannian geometry, to explain concepts that have logical and historical priority. Hopefully students of geometry whose experiences parallel my own will benefit from this approach, while readers for whom these efforts are unnecessary will forgive certain sections which, like this preface, may seem a bit verbose. Regardless, I hope you will find this subject as interesting and rewarding as I have.

# Acknowledgments

I must first thank my advisor, Gary Jensen, who basically taught me everything I know about geometry. The amount of time and energy you invested in me over the past four years is staggering. Your impressive work ethic and general approach to thinking about mathematics have been an invaluable compass in my own attempt to become a mathematician, and I thank you for being a friend and personal mentor. More generally, the entire WU math department has been overwhelmingly supportive, particularly the very talented faculty: thank you for your ceaseless efforts to design exciting courses and share these beautiful ideas. I also thank the dissertation committee for their help and patience along the way.

The greatest debt of gratitude is owed to my parents, James and Deborah. Thank you for encouraging and supporting my academic efforts since childhood. Your integrity and passion in pursuit of your own dreams has inspired me to follow mine, your constant and unconditional love and support have enabled me to do so. I love you both very much and I can never thank you enough for everything you've done. Thank you Gabe for being a great brother and basically the most genuine person I know: your determination and courageous approach to life is literally awesome to me, and much of my own motivation derives from the pleasure of imitating your example. I can't even guess at the tremendous help I've received from the rest of my St. Louis family, so I'll just mention that Mimi, Bin, and all my aunts and uncles are the best. Clayton, Jesse, and Jeff may as well be included here since they are essentially brothers to me.

Friends and colleagues are uncountable and I'll only name a few here: thank you Xiao for a rather wonderful graduate life, adventuring, TeX help, da qiu, and general drama. Quite similarly, thank you Tori for climbing, baked goods, understanding, and various unspeakables. Scotty and Nic were extra nice for sticking around with me after everyone else left, also particularly true for Fatima (MD). Josh, Andy, Drew, and Sara are in a similar category, as well as the massive Chinese contingency. Special thanks to Larry for being a great guy, climber, psychotic - Aaron likewise. Brian, Tim, Raj, Brad, Meghana, Jonathan, Sooraj and the PTs are some of the best friends anyone could hope for; Jeff Blanchard, David Opela, and Amei are personal heroes. Instead of listing more names, let me mention that a visiting faculty once commented that the camaraderie among the graduate students in our department was like nothing he'd ever seen. The past years at the WU math department have been the happiest time of my life and I think this largely why. I hope future students never loose that feeling of community, it makes this department a very special place in my opinion.

# Contents

# List of Figures

# Chapter 1

# Introduction to Surface Theory

## 1.1 Immersions into the Space forms

A *surface* is a real 2-dimensional smooth manifold. We first consider immersions of a surface $M^2$ into 3-dimensional Euclidean space, that is, smooth maps $\mathbf{x} : M^2 \to \mathbf{R}^3$ whose Jacobian $d\mathbf{x}$ has full rank everywhere. This is essentially the starting point of Gauss' study of differential geometry [8], which is often regarded as the birth of the subject. We will implicitly identify $M^2$ with its image $\mathbf{x}(M^2) \subset \mathbf{R}^3$, or think of $M^2$ as the parameter domain for a surface in space. The Euclidean metric on $\mathbf{R}^3$ pulls back via $\mathbf{x}$ to a metric on $M^2$ called the *induced metric*, or in Gauss' terminology, the *first fundamental form* $I = ds^2 := d\mathbf{x} \cdot d\mathbf{x}$, where "$\cdot$" denotes the usual Euclidean inner product and the differential of $\mathbf{x}$ is taken component-wise. When Gauss began his study, the curvature of a smooth curve $\boldsymbol{\gamma} : \mathbf{R} \to \mathbf{R}^3$ was already a well-established notion (essentially known to the Greek geometers in a non-differential form), but it was unclear how to generalize this notion to surfaces. Given a point $p \in M^2$, we can find an open neighborhood $U \in M^2$ of $p$ on which there exists a smooth unit vector field $\mathbf{n} : U \to S^2 \subset \mathbf{R}^3$ which is normal to the surface at every point, called the *Gauss map*. Gauss recognized that the curvature of the surface should be very much related to the extent to which $\mathbf{n}$ is changing. Using parallel translation to identify the tangent spaces $T_pM^2$ and $T_{\mathbf{n}(p)}S^2$, we can regard the differential $-d\mathbf{n} : T_pM^2 \to T_{\mathbf{n}(p)}S^2$ as an endomorphism of $T_pM^2$, called the *shape operator*, and define its associated bilinear form $II := -d\mathbf{x} \cdot d\mathbf{n}$, called the *second fundamental form.* If $\mathcal{S}$ is the matrix representation of $-d\mathbf{n}$ with respect to a local frame for $TU$, we can apply the elementary symmetric functions to this form and obtain (basis independent) notions of curvature:[1]

$$K := det\,\mathcal{S} \qquad\qquad H := \frac{1}{2}\,tr\,\mathcal{S}$$

$K$ is called the *Gaussian curvature*, so named for the great theorem of Gauss demonstrating that $K$ depends only on the induced metric. Regarding the metric as an "internal" feature of the geometry of the surface, we say that $K$ is an *intrinsic* invariant of $M^2$ (unlike $H$, the *Mean curvature*, which depends on how $M^2$ is immersed in space, and is therefore considered *extrinsic*).

The first and second fundamental forms $I$ and $II$ on $M^2$ satisfy a system of PDE called the *Gauss and Codazzi equations*, which uniquely determinethe immersion $\mathbf{x} : M^2 \to \mathbf{R}^3$ up to a rigid motion, in the sense that:

**Theorem 1.** (Bonnet's Uniqueness Theorem) *Let $\mathbf{x}, \tilde{\mathbf{x}} : M^2 \to \mathbf{R}^3$ be two immersions of an oriented and connected surface $M^2$. Then $\mathbf{x}$ and $\tilde{\mathbf{x}}$ are congruent if and only if $I = \tilde{I}$ and $II = \widetilde{II}$. More precisely, $\tilde{\mathbf{x}} = A\mathbf{x} + \mathbf{v}$ for some fixed rotation $A \in O_3(\mathbf{R})$ and translation $\mathbf{v} \in \mathbf{R}^3$ if and only if there exist unit normal vector fields $\mathbf{n}$, $\tilde{\mathbf{n}}$ along $\mathbf{x}$, $\tilde{\mathbf{x}}$ respectively, with respect to which the second fundamental forms agree.*

[1] Alternatively, given a unit tangent vector $\mathbf{v}$ at the point $p$, consider the intersection of $M^2$ with the unique plane containing $p$, $\mathbf{v}$ and $\mathbf{n}_p$. This intersection is a curve with curvature $\kappa(\mathbf{v})$ at $p$. Regarding $\kappa$ as a function on $S^1$ and letting $\kappa_1$ and $\kappa_2$ denote its maximum and minimum, $K(p) = \kappa_1\kappa_2$ and $H(p) = \frac{1}{2}(\kappa_1 + \kappa_2)$. These $\kappa_i$ are exactly the eigenvalues of the shape operator, called the *principal curvatures* by Euler.

This is really half of *Bonnet's Theorem*, the more general form of which is also referred to as the *Fundamental Theorem of Hypersurfaces*. The other half states that given *any* bilinear forms $I$, $II$ on $M^2$ satisfying the Gauss and Codazzi equations, there exists an immersion $\mathbf{x} : M^2 \to \mathbf{R}^3$ with $I$, $II$ as its fundamental forms. Consequently, we may think of these PDE as "integrability" conditions for the immersion. We can state these equations by reformulating Gauss' theory in the language of Cartan's method of moving frames, after which Bonnet's theorem can be obtained from the Cartan-Darboux theorem. For a more general and complete account with proofs of the theorems that follow, see Jensen's book on moving frames [12].

Given local coordinates $(u, v)$ on $M^2$, the vectors $\mathbf{x}_u, \mathbf{x}_v$ span the tangent plane to the surface at each point, and can be used to build an orthonormal frame field $\{e_1, e_2, e_3\}$ for $\mathbf{R}^3$ along $M^2$ by the Gram-Schmidt process. Note that $e_3$ is now the Gauss map $\mathbf{n}$. The frame vectors form columns of a matrix-valued map $e = (e_1, e_2, e_3) : M^2 \to O_3(\mathbf{R})$ into the orthogonal group. The method of moving frames then says that differential invariants of the surface $M^2$ can be found by examining the pull-back to $M^2$ of the $\mathfrak{o}_3(\mathbf{R})$-valued Maurer-Cartan 1-form $\omega$ on $O_3(\mathbf{R})$ via the frame $e$. This form can be written $e^*\omega = e^{-1}de$, which by the standard abuse of notation we will simply denote $\omega$:

$$\omega \;=\; e^{-1}de \;=\; \begin{pmatrix} \omega_1^1 & \omega_2^1 & \omega_3^1 \\ \omega_1^2 & \omega_2^2 & \omega_3^2 \\ \omega_1^3 & \omega_2^3 & \omega_3^3 \end{pmatrix} \;=\; \begin{pmatrix} 0 & \omega_2^1 & \omega_3^1 \\ -\omega_2^1 & 0 & \omega_3^2 \\ -\omega_3^1 & -\omega_3^2 & 0 \end{pmatrix}$$

So we have 1-forms $\omega_j^i = e_i \cdot de_j$ which satisfy the structure equations $d\omega_j^i = -\omega_k^i \wedge \omega_j^k$. We also define 1-forms $\omega^i$ by $\omega^i = d\mathbf{x} \cdot e_i$, which satisfy[2] $d\omega^i = -\omega_j^i \wedge \omega^j$. Note that $\omega^1, \omega^2$ constitute an *orthonormal coframe* on $M^2$ (that is, 1-forms putting the metric

[2]Had we included $\mathbf{x}$ itself as the first column of the frame, as is customary, and regarded $e$ as a map into the Euclidean group $\mathbf{R}^3 \rtimes O_3(\mathbf{R})$, both equations could be succinctly written in matrix form as $d\omega = -\omega \wedge \omega$, the hallmark relation uniquely satisfied by the Maurer-Cartan 1-form $\omega$.

in the form $I = \omega^1\omega^1 + \omega^2\omega^2$), since $e_3$ being normal to the surface is equivalent to the condition that $\omega^3 \equiv 0$ on $M^2$. By differentiating this condition, the structure equations imply that $\omega_1^3 \wedge \omega^1 + \omega_2^3 \wedge \omega^2 = 0$. Since $\omega^1$, $\omega^2$ form a basis for the 1-forms on $M^2$, this last equation together with Cartan's lemma tells us that the $\omega_i^3$ can be written as a linear combination of the $\omega^i$ with *symmetric* coefficient functions:

$$\omega_i^3 = h_{ij}\,\omega^j, \quad \text{where} \quad h_{ij} = h_{j\,i}$$

But $(h_{ij})$ is exactly the matrix $\mathcal{S}$ of $de_3$ with respect to the basis $\{e_1, e_2\}$. That this matrix is symmetric says that the second fundamental form is symmetric $II = \omega^1\omega_1^3 + \omega^2\omega_2^3 = h_{ij}\,\omega^i\omega^j$ (which is the case for general hypersurfaces in a Riemannian manifold). Again by the structure equations we have $d\omega_2^1 = \omega_1^3 \wedge \omega_2^3 = (h_{11}h_{22} + h_{12}^2)\,\omega^1 \wedge \omega^2$, which is called the *Gauss equation*:

$$d\omega_2^1 = K\omega^1 \wedge \omega^2$$

Differentiating $\omega_i^3 = h_{ij}\,\omega^j$ and applying Cartan's lemma, we obtain smooth functions $h_{ijk}$, totally symmetric is all three indices, satisfying the *Codazzi equations*:

$$dh_{ij} - h_{kj}\,\omega_i^k - h_{ik}\,\omega_j^k = h_{ij\,k}\,\omega^k$$

The $h_{ijk}$ are symmetric in their first two indices because the $h_{ij}$ are, so the Codazzi equations are sometimes taken as the statement that $h_{ijk}$ are also symmetric in the last two indices.

The second half of Bonnet's theorem can now be stated as follows. We assume here that the surface $M^2$ is simply connected, which makes the theorem a local result:

**Theorem 2.** (Bonnet's Existence Theorem) *Let $I$ and $II$ be two symmetric bilinear forms on $M^2$, with $I$ positive definite, satisfying the Gauss and Codazzi equations in the following sense: Given any orthonormal coframe field $\{\omega^1, \omega^2\}$ for $I$ (that is,* *linearly independent 1-forms such that* $I = \omega^1\omega^1 + \omega^2\omega^2$*), suppose that the 1-forms* $\omega^i_j = -\omega^j_i$ *such that* $d\omega^i = -\omega^i_j \wedge \omega^j$ *and the functions* $h_{ij}$ *such that* $II = h_{ij}\,\omega^i\omega^j$ *satisfy the equations*

$$\begin{cases} d\omega^1_2 = (h_{11}h_{22} + h^2_{12})\,\omega^1 \wedge \omega^2, \\ dh_{ij} - h_{kj}\,\omega^k_i - h_{ik}\,\omega^k_j = h_{ij\,k}\,\omega^k, \quad \text{for some functions } h_{ij\,k} = h_{ikj}. \end{cases}$$

*Then there exists an immersion* $\mathbf{x} : M^2 \to \mathbf{R}^3$ *with Gauss map* $\mathbf{n} : M^2 \to S^2$ *such that* $I = d\mathbf{x}\cdot d\mathbf{x}$ *and* $II = -d\mathbf{x}\cdot d\mathbf{n}$*.*

As mentioned above, Bonnet's theorem can be generalized to higher dimensions as well as to non-flat ambient spaces. This is almost immediate for the *space forms* (the manifolds of constant sectional curvature): Denote by $S^3_0$ Euclidean 3-space $\mathbf{R}^3$, $S^3_1$ the sphere of radius 1 in $\mathbf{R}^4$, $S^3_{-1}$ (one sheet of) the hyperboloid of radius $-1$ in Minkowski 4-space $\mathbf{R}^{1,3}$ (recall that $\mathbf{R}^{1,3}$ is $\mathbf{R}^4$ equipped with the pseudo-metric $\langle \mathbf{x}, \mathbf{y}\rangle := \mathbf{x}^T L \mathbf{y}$ where $L = diag(-1,1,1,1)$), and $G_\epsilon = \mathit{Isom}(S^3_\epsilon)$ the group of orientation-preserving isometries on $S^3_\epsilon$:

$$S^3_\epsilon = \begin{cases} S^3 = \{\mathbf{x} \in \mathbf{R}^4 \mid \|\mathbf{x}\| = 1\} & \epsilon = 1 \\ \mathbf{R}^3 & \epsilon = 0 \\ H^3 = \{\mathbf{x} \in \mathbf{R}^{1,3} \mid \|\mathbf{x}\| = -1\} & \epsilon = -1 \end{cases}$$

$$G_\epsilon = \begin{cases} SO_4(\mathbf{R}) = \left\{O \in SL_3(\mathbf{R}) \mid O^T O = I\right\} & \epsilon = 1 \\ SE_3(\mathbf{R}) = \mathbf{R}^3 \rtimes SO_3(\mathbf{R}) & \epsilon = 0 \\ SO_{1,3}(\mathbf{R}) = \left\{O \in SL_3(\mathbf{R}) \mid O^T L O = L\right\} & \epsilon = -1 \end{cases}$$

$G_\epsilon$ acts transitively on $S^3_\epsilon$ with isotropy subgroup $SO_3(\mathbf{R}) \subset G_\epsilon$ at the points $\mathbf{0} \in S^3_0$ and $(1,0,0,0) \in S^3_1, S^3_{-1}$, so that $S^3_\epsilon \simeq G_\epsilon/SO_3(\mathbf{R})$. One computes that the Maurer-Cartan 1-form $\omega$ on $G_\epsilon$ can be written

$$\omega = \begin{pmatrix} 0 & -\epsilon\omega^i \\ \omega^i & \omega^j_k \end{pmatrix} \quad \text{where } (\omega^j_k) \in \mathfrak{so}_3(\mathbf{R})$$

and the structure equations $d\omega = -\omega \wedge \omega$ then show[3] that $S^3_\epsilon$ has constant sectional curvature $\epsilon$, so these are indeed (models of) the 3-dimensional space forms.

Now suppose that $\mathbf{x} : M^2 \to S^3_\epsilon$ is an immersion. A *frame* along $\mathbf{x}$ is a lift, that is, a map $e : M^2 \to G_\epsilon$ such that $\pi \circ e = \mathbf{x}$, where $\pi : G_\epsilon \to S^3_\epsilon$ is natural projection. With the usual abuse of notation let $\omega = e^{-1}de$ denote the pullback of the Maurer-Cartan form by $e$. Defining functions $h_{ij}$ as before, the structure equations imply that $d\omega^1_2 = \omega^3_1 \wedge \omega^3_2 + \epsilon\, \omega^1 \wedge \omega^2 = (h_{11}h_{22} + h^2_{12} + \epsilon)\, \omega^1 \wedge \omega^2$, which is the Gauss equation:

$$d\omega^1_2 = (K + \epsilon)\, \omega^1 \wedge \omega^2$$

Bonnet's Theorem can now be restated almost verbatim for immersions into the space forms $\mathbf{x} : M^2 \to S^3_\epsilon$, where "$\mathbf{x}$ and $\tilde{\mathbf{x}}$ are congruent" means that there is an element $A \in G_\epsilon$ such that $\mathbf{x} = A\tilde{\mathbf{x}}$, the Gauss equation reads as above, and the Codazzi equations remain the same.

## 1.2 The induced complex structure

The term "Riemann surface" appears at first to have a peculiar double usage. On the one hand, Riemann introduced the curvature tensor $R^i_{jkl}$ which captures the most general concept of (metric) curvature, including the Gaussian curvature as a special case. Any manifold with a metric $(M^n, g)$ is called a *Riemannian manifold* on account of this accomplishment, and "Riemann surface" naturally denotes the 2-dimensional case. On the other hand, Riemann was also instrumental in the development of classical function theory, particularly his observation that multi-valued holomorphic functions on the complex plane may become single-valued when regarded

[3]Namely: $\Omega^i_j := d\omega^i_j + \omega^i_k \wedge \omega^k_j = \epsilon\omega^i \wedge \omega^j = R^i_{jkl}\omega^k \wedge \omega^l \implies R^i_{jkl} = \epsilon(\delta^i_k\delta^j_l - \delta^i_l\delta^j_k)$. $\square$

as functions on some "larger" domain. From this he pioneered the highly influential philosophy that the properties of a holomorphic function can be encoded in the geometry of this underlying domain, called the associated *Riemann surface*. This surface is a 1-dimensional complex manifold, or *complex curve*, and consequently "Riemann surface" also refers to any such curve.

Miraculously, there is no ambiguity in these two uses of the term! To be precise, any *conformal structure* on a surface $M^2$ (that is, a metric defined up to multiplication by a smooth positive function) gives rise to a complex structure on $M^2$, and conversely. This is essentially the content of the Korn-Lichtenstein theorem, of which we will be making constant implicit use in the sequel. Recall that a Riemannian manifold $(M^n, g)$ is called *flat* if it is locally isometric to Euclidean space, i.e. at any point on $M^n$, there exists coordinates $(U, x^i)$ such that the metric can be written as $g = \sum (dx^i)^2$ on $U$, and thus $(U, g\,|_U)$ is isometric to $(\mathbf{R}^n, \sum (dx^i)^2)$. As noted above, a curve $C \subset \mathbf{R}^n$ has no intrinsic geometry since it must be flat, parametrization $\boldsymbol{\gamma} : \mathbf{R} \to C$ by arclength $s$ providing an isometry of $(\mathbf{R}, dx)$ with $(C, ds)$. Having proved that flatness is equivalent to the vanishing of the curvature tensor, Riemann considered a more general notion of flatness:

**Definition 3.** A Riemannian manifold $(M^n, g)$ is called *conformally flat* if it is locally conformal[4] to Euclidean space, i.e. at any point on $M^n$, there exists coordinates $(U, x^i)$ such that the metric can be written as $g = e^{2u} \sum (dx^i)^2$ on $U$ for some smooth function $u : U \to \mathbf{R}$.

Coordinates putting the metric in this form are called *isothermal coordinates* and the function $u$ is called the *conformal factor* relative to these coordinates. As we shall

[4]More generally, given two Riemannian manifolds $(M, g)$, $(\widetilde{M}, \tilde{g})$, a map $\phi : M \to \widetilde{M}$ is said to be *conformal* if $\phi^* \tilde{g} = e^{2u} g$ for some smooth function $u : M \to \mathbf{R}$, and $g$ is said to be *conformally related* to $\tilde{g}$. When $(M, g) = (\widetilde{M}, \tilde{g})$, this is exactly the condition that $\phi$ preserves angles.

see, although a surface need not be flat ($K \not\equiv 0$), it must be *conformally* flat. First recall the following definitions:

**Definition 4.** Let $M$ be a real manifold of even dimension.

1.) An *almost complex structure* on $M$ is a smooth (1,1)-type tensor field $J$ such at each point $p \in M$, $J_p : T_pM \to T_pM$ satisfies $J_p^2 = -Id_{T_pM}$.

2.) A *complex structure*[5] on $M$ is a cover of $M$ by complex coordinate charts $(U_\alpha, z_\alpha^i)$, $z_\alpha^i : U \to \mathbf{C}$ such that the transition functions $z_\alpha^i \circ z_\beta^{j\,-1} : \mathbf{C} \to \mathbf{C}$ are holomorphic.

An almost complex structure provides an identification of $T_pM$ with $\mathbf{C}^n$, defining complex scalar multiplication on $T_pM$ by $(a+bi)X := aX + bJX$. A complex structure clearly defines an almost complex structure by pulling back multiplication by $i$ via the isomorphism $dz_p : T_pM \to \mathbf{C}^n$ (that is, $J_p := dz_p^{-1} \circ i \circ dz_p$), but an almost complex structure does not necessarily define a complex structure, unless it satisfies certain integrability conditions (in which case it is called *integrable*).

Now consider the special case of an oriented surface $M = M^2$ with a metric $I$. Let $\omega^1$, $\omega^2$ be any local orthonormal coframe for $I = \sum(\omega^i)^2$, and define a complex 1-form $\phi \in \bigwedge^1(U; \mathbf{C})$ by $\phi := \omega^1 + i\omega^2$. This is an almost complex structure, since $\phi_p : T_p(M^2) \to \mathbf{C}$ is an isomorphism at each point $p \in M^2$, and we can define $J_p := \phi_p^{-1} \circ i \circ \phi_p$ as above. Note that $J$ depends neither on the choice of coframe (since if $(\tilde{\omega}^i)$ is another coframe for $I$, it will be related to the original by $(\tilde{\omega}^i) = A(\omega^i)$, with $A : U \to SO_2(\mathbf{R}) \simeq U(1) \subset \mathbf{C}$, so that $\tilde{\phi} = A\phi$ and $\tilde{J}_p = \tilde{\phi}_p^{-1} \circ i \circ \tilde{\phi}_p = \phi_p^{-1} \circ A^{-1} i A \circ \phi_p = \phi_p^{-1} \circ i \circ \phi_p = J_p$) nor the choice of "conformal representative" of the metric (in the sense that if $\tilde{I} = e^{2u} I$ with $u : U \to \mathbf{R}$ is another representative of $I$, then $(\tilde{\omega}^i) = e^u(\omega^i)$ is a coframe field for $\tilde{I}$, so by the same calculation $\tilde{J}_p = J_p$).

[5] A *complex manifold* is just a manifold with a complex structure.

But the remarkable fact is that *any* almost complex structure on a surface must be integrable:

**Theorem 5.** *(Korn-Lichtenstein) Given a surface with a metric $(M^2, I)$, define the complex 1-form $\phi := \omega^1 + i\omega^2$. Then at every point there exist local complex functions $f, z : U \to \mathbf{C}$ such that $\phi = f\,dz$ on $U$, and the complex coordinate charts $(U_\alpha, z_\alpha)$ constitute a complex structure. Conversely, any complex structure $(U_\alpha, z_\alpha)$ on a surface $M^2$ can be obtained from some metric $I$ on $M^2$ in this way.*

The complex structure whose existence is asserted by the theorem is call the *induced complex structure*. Notice that $I = \phi\bar{\phi} = |f|^2\,|dz|^2 = |f|^2\,(dx^2+dy^2)$ where $z = x+iy$, so $(x, y)$ are indeed isothermal coordinates and thus any surface is conformally flat.

With this extra structure in place, we again consider an immersion $\mathbf{x} : M^2 \to \mathbf{R}^3$, which induces a metric $ds^2 = d\mathbf{x} \cdot d\mathbf{x}$, which in turn induces a complex structure $z$. We would like to express the invariants of the immersion in terms of the complex variable $z$ and reformulate the Gauss and Codazzi equations appropriately. As noted above, rotating a coframe $(\omega^i)$ by an angle $\theta$ has the effect of multiplying $\phi$ by $e^{i\theta}$, thereby multiplying $f$ by $e^{i\theta}$. Thus $(\tilde{\omega}^i) = e^{-i\arg(f)}(\omega^i)$ puts $\tilde{\omega}^1 + i\tilde{\omega}^2 = e^u dz$ where $u : U \to \mathbf{R}$ is the conformal factor. A frame $e$ along $\mathbf{x}$ such that $\omega^1 + i\omega^2 = e^u dz$ is said to be *adapted* to the complex structure, and this frame is unique.

**Definition 6.** Let $e$ be the frame adapted to the local complex coordinate $(U, z)$ and $h_{ij}$ the real functions such that $\omega_i^3 = h_{ij}\,\omega^j$. The complex function $h : U \to \mathbf{C}$ defined by $h := \frac{1}{2}(h_{11} - h_{22}) - ih_{12}$ is called the *Hopf invariant* of $\mathbf{x}$ relative to coordinate $z$.

Notice that $|h|^2 = H^2 - K$, which we will use to rewrite the Gauss and Codazzi equations in terms of the Hopf invariant $h$, the conformal factor $u$, and the mean curvature $H$. First observe that $\omega^1 + i\omega^2$ can be differentiated in two ways[6]

[6]Many of the calculations in this sections rely on the unique decomposition of tensors into *bide-*

$$d(\omega^1 + i\omega^2) \;=\; -\omega^1_2 \wedge \omega^2 - i\omega^2_1 \wedge \omega^2 \;=\; i\omega^1_2 \wedge (\omega^1 + i\omega^2) \;=\; i\omega^1_2 \wedge e^u dz$$

$$d(\omega^1 + i\omega^2) \;=\; d(e^u dz) \;=\; u_{\bar z}\, d\bar z \wedge e^u dz$$

which shows that $\omega^1_2 \wedge e^u dz = -iu_{\bar z}\, d\bar z \wedge e^u dz$, and since $\omega^1_2$ is real, $\omega^1_2 = Re(-iu_{\bar z}\, d\bar z) = -iu_{\bar z}\, d\bar z + iu_z\, dz$. Differentiating this we have $d\omega^1_2 = -2iu_{z\bar z}\, dz \wedge d\bar z = -4e^{-2u} u_{z\bar z}\, \omega^1 \wedge \omega^2$, since the area form is $dA = \omega^1 \wedge \omega^2 = \frac{i}{2}(\omega^1 + i\omega^2) \wedge (\omega^1 - i\omega^2) = \frac{i}{2} e^{2u} dz \wedge d\bar z$. By the original Gauss equation $d\omega^1_2 = K\omega^1 \wedge \omega^2$, we conclude that $K = -4e^{-2u} u_{z\bar z}$. Combining this with $|h|^2 = H^2 - K$ we can restate the Gauss equation as:

$$-4e^{-2u} u_{z\bar z} \;=\; H^2 - |h|^2$$

Next use $\omega^3_i = h_{ij}\, \omega^j$ to write $\omega^3_1 - i\omega^3_2 = h\,(\omega^1 + i\omega^2) + H(\omega^1 - i\omega^2) = h\, e^u dz + He^u d\bar z$. Then $\omega^3_1 + i\omega^3_2$ can be differentiated in two ways:

$$\begin{aligned}
d(\omega^3_1 + i\omega^3_2) \;&=\; -\omega^3_2 \wedge \omega^2_1 - i\omega^3_1 \wedge \omega^1_2 \;=\; i\omega^1_2 \wedge (\omega^3_1 - i\omega^3_2) \\
&=\; (u_{\bar z}\, d\bar z - u_z\, dz) \wedge (h\, e^u dz + He^u d\bar z) \\
&=\; e^u(-h\, u_{\bar z} - Hu_z)\, dz \wedge d\bar z \\
d(\omega^3_1 + i\omega^3_2) \;&=\; d(h\, e^u dz + He^u d\bar z) \;=\; (h_{\bar z}\, e^u + h\, u_{\bar z})\, dz \wedge d\bar z + (H_z e^u + Hu_z)\, d\bar z \wedge dz \\
&=\; e^u(h_{\bar z} + h\, u_{\bar z} - H_z - Hu_z)\, dz \wedge d\bar z
\end{aligned}$$

which when compared yield the complex Codazzi equations:

$$h_{\bar z} + 2h\, u_{\bar z} \;=\; H_z$$

An immediate application of the complex structure equations is the following characterization of *CMC immersions*, that is, immersions of *constant mean curvature*,

*grees*, which amounts to using the local $C^\infty(U)$-basis for the tensor algebra $\mathcal{T}^*U$ generated by $\partial_z$, $\partial_{\bar z}$, $dz$, and $d\bar z$.

$H \equiv const.$ Decompose the second fundamental form $II$ into bidegrees $II = II^{2,0} + II^{1,1} + II^{0,2}$ by choosing an adapted frame $e$ so that

$$\begin{aligned} II &= \omega^1\omega_1^3 + \omega^2\omega_2^3 \;=\; Re[(\omega^1 + i\omega^2)(\omega_1^3 - i\omega_2^3)] \;=\; Re[(e^u dz)(h\, e^u dz + He^u d\bar{z})] \\ &= \frac{1}{2}[\, e^{2u} h\, dzdz \;+\; 2He^u dzd\bar{z} \;+\; e^{2u} h\, d\bar{z}d\bar{z}] \\ &= \qquad II^{2,0} \quad + \quad II^{1,1} \quad + \quad II^{0,2} \end{aligned}$$

The (2,0)-bidegree $II^{2,0} = \frac{1}{2} e^{2u} h\, dzdz$ is called the *Hopf differential* of $\mathbf{x}$.

**Proposition 7.** *An immersion* $\mathbf{x}$ *is CMC if and only if* $II^{2,0}$ *is holomorphic.*

*Proof.* $II^{2,0}$ is holomorphic on $(U, z)$ $\Leftrightarrow$ $0 = (e^{2u} h\,)_{\bar{z}} = e^{2u}(h_{\bar{z}} + 2h\, u_{\bar{z}}) = e^{2u} H_z$ $\Leftrightarrow$ $H_z = 0$ $\Leftrightarrow$ $H$ is constant on $(U, z)$. Applying the argument to a cover of coordinate charts, continuity implies that $H$ is constant on all of $M^2$. $\square$

For completeness, we conclude this section by restating Bonnet's Theorem for immersions into the space forms.

**Theorem 8.** *Given functions* $H, u : M^2 \to \mathbf{R}$ *and* $h : M^2 \to \mathbf{C}$ *on a simply connected Riemann surface* $M^2$ *satisfying the structure equations*

$$\begin{cases} -4e^{-2u} u_{z\bar{z}} \;=\; H^2 - |h\,|^2 + \epsilon & \text{(Gauss equation)} \\ h_{\bar{z}} + 2h\, u_{\bar{z}} \;=\; H_z & \text{(Codazzi equation)} \end{cases}$$

*there exists a conformal immersion* $x : M^2 \to S^3_\epsilon$ *with invariants* $H$, $u$, *and* $h$. *This immersion is unique up to rigid motion, in the sense that two conformal immersions* $x, \tilde{x} : M^2 \to S^3_\epsilon$ *whose mean curvature, conformal factor, and Hopf invariant agree with respect to every local complex coordinate, there exists* $A \in G_\epsilon$ *such that* $\tilde{x} = Ax$.

## 1.3 Minimal surfaces

A curve $\gamma : \mathbf{R} \to M$ is called a *geodesic* if it locally minimizes length. A submanifold $x : N \to M$ is called *minimal* if it locally minimizes volume. Minimal submanifolds are certainly a very natural set of objects in Riemannian geometry, and a huge number of the classical geometers seem to have spent at least a little time considering them. We will be interested in the simplest possible case, minimal surfaces in Euclidean 3-space $\mathbf{x} : M^2 \to \mathbf{R}^3$, where implications of the induced complex structure on $M^2$ and special features of dimension three combine to yield a stunningly rich theory.

Given a closed curve $C$ in $\mathbf{R}^3$, one can try to find a surface $M^2$ which minimizes area among all surfaces with the given curve as their boundary $\partial M^2 = C$. This is called the *Plateau problem*, after the nineteenth-century Belgian physicist Plateau who approached the problem using the following physical picture: He imagined $C$ to be a loop of rigid wire, which is then dipped into a solution of soapy water. As the wire is removed from the solution, a soap bubble will form, and the surface tension in the bubble will tend to minimize its total area, providing an approximation to the desired minimal surface $M^2$ with $\partial M^2 = C$. This will be a helpful picture to keep in mind during what follows, but we will not be concerned with the Plateau problem here. We want instead to consider minimal surfaces *without boundary*, defined by the property that for any point $p \in M^2$ on the surface, there exists a neighborhood $U \subset M^2$ such that the closure $\bar{U}$ solves the Plateau problem for the curve $C = \partial U$ (that is, $M^2$ *locally* minimizes surface area).

Lagrange seems to have been the first to successfully study minimal surfaces in this later sense. In his 1762 treatise [13], he used the Calculus of Variations to obtain a quasi-linear elliptic PDE (the *minimal surface equation*, or *MSE*) which must be satisfied by a function $f(x, y)$ on $\mathbf{R}^2$ whose graph is a minimal surface. He noted that it is satisfied by the plane, which was the only known example until 1776, when Meusnier [15] rewrote this equation in a simplified form and obtained the two solutions of great importance: the Catenoid ($z = f(x, y) = \cosh^{-1}\sqrt{x^2 + y^2}$) and the Helicoid ($z = f(x, y) = \tan^{-1}\frac{x}{y}$). The Catenoid is essentially the only minimal surface of revolution, and the Helicoid is essentially the only ruled minimal surface.

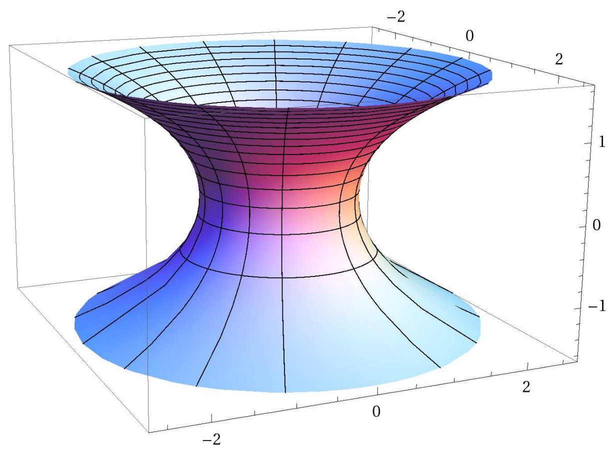

**Fig.1.1** *Catenoid*

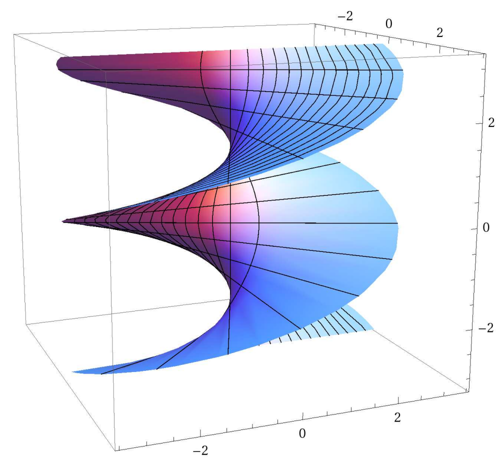

**Fig.1.2** *Helicoid*

We shall not deal with the MSE directly, but the ideas that lead to its derivation are quite essential. In the language of the Calculus of Variations, minimality of $\mathbf{x} : M^2 \to \mathbf{R}^3$ is equivalent to the condition that $\mathbf{x}$ is a critical point for the *area functional* $A$, that is, the map that assigns to an immersion $\mathbf{x}$ its surface area:

$$A(\mathbf{x}) = \int_{M^2} dA$$

where "$dA$" is the area form of the metric on $M^2$ induced by $\mathbf{x}$. In order to make sense of the derivative of $A$, we define a *variation* of $\mathbf{x}$ to be a 1-parameter family of immersions $\mathbf{x}_t$ of the form $\mathbf{x}_t = \mathbf{x} + tfe_3$, where $f : M^2 \to \mathbf{R}$ is a smooth function with compact support and $e_3$ is the Gauss map of $\mathbf{x}$. We can differentiate $A(\mathbf{x}_t) : \mathbf{R} \to \mathbf{R}$ as a function of $t$ and evaluate the result at $t = 0$, and $\mathbf{x}$ is called *critical* for $A$ if this is zero for all variations of $\mathbf{x}$. A standard calculation shows that

$$\frac{d}{dt}A(\mathbf{x}_t) = -2\int_{M^2} fHdA$$

so $\mathbf{x}$ is critical for $A$ if and only if $\int_{M^2} fHdA = 0$ for *any* compactly supported $f$. The Fundamental Lemma of Variations then proves the basic characterization of minimal surfaces[7], first observed by Meusnier:

**Theorem 9.** *(Lagrange-Meusnier) An immersion* $\mathbf{x} : M^2 \to \mathbf{R}^3$ *is minimal if and only if its mean curvature vanishes identically* $H \equiv 0$.

This fact instantly provides a wealth of geometric insight: The principle curvatures of a minimal surface $M^2$ must be equal and opposite at every point ($K \leq 0$), so that $M^2$ is locally saddle-shaped, except at umbilic[8] points (as one might expect examining Plateau's soap bubbles). We also note that if $\mathcal{S} = (h_{ij})$ is the matrix of the shape operator, then $\mathcal{S}^2$ is the matrix of the *third fundamental form*, $\mathit{III} := de_3 \cdot de_3$, the pullback of the round metric on $S^2$ by the Gauss map $e_3 : M^2 \to S^2$. By Cayley-Hamilton, $\mathcal{S}^2 - tr(\mathcal{S})\mathcal{S} + det(\mathcal{S})I = 0$, so the fundamental forms are related according to $\mathit{III} - 2H\,\mathit{II} + K\,I = 0$. If $\mathbf{x}$ is minimal, we have $\mathit{III} = -K\,I$, so the Gauss map is conformal.[9] This also occurs when $\mathbf{x}$ is *totally umbilic* ($\mathit{II} = \lambda I$ for some $\lambda : M^2 \to \mathbf{R}$), since then $\mathit{III} = (2H\lambda - K)\,I$. Excluding this case, we can obtain another condition equivalent to minimality:

**Proposition 10.** *Suppose* $\mathbf{x} : M^2 \to \mathbf{R}^3$ *is not totally umbilic. Then* $\mathbf{x}$ *is minimal if and only if its Gauss map is conformal.*

While this fact is indeed useful, the most crucial property for our purposes is the

[7] In fact, this characterization holds for arbitrary dimension and codimension.

[8] A point $p \in M^2$ is called *umbilic* if the principal curvatures at $p$ are equal $\kappa_1 = \kappa_2$, if and only if the matrix $(h_{ij})|_p = \lambda I$ for some $\lambda \in \mathbf{R}$, if and only if the Hopf differential vanishes $\mathit{II}^{2,0} = 0$.

[9] This also shows that the metric $-K\,I$ on $M^2$ has constant curvature 1 when $\mathbf{x}$ is minimal, since the round metric has constant curvature 1. This condition is also equivalent to minimality (Ricci).

intimate connection between minimal immersions of $M^2$ and the induced complex structure, established as follows

**Proposition 11.** *An immersion* $\mathbf{x} : M^2 \to \mathbf{R}^3$ *is minimal if and only if its coordinate functions are harmonic.*

*Proof.* The coordinate functions of $\mathbf{x} = (x_1, x_2, x_3)$ are harmonic if and only if $\mathbf{x}_{z\bar{z}} = 0$. If $e$ is an adapted frame along $\mathbf{x}$, then $\mathbf{x}_z = \frac{1}{2}\, e^u(e_1 - ie_2)$, so $d\mathbf{x}_z = \frac{1}{2}\, d(e^u)(e_1 - ie_2) + \frac{1}{2}\, e^u d(e_1 - ie_2)$. First, we use the identities $-i\omega_2^1 = u_z\, dz - u_{\bar{z}}\, d\bar{z}$ and $\omega_1^3 - i\omega_2^3 = h\, e^u dz + He^u d\bar{z}$ to compute

$$\begin{aligned} d(e_1 - ie_2) \;&=\; \omega_1^2 e_2 + \omega_1^3 e_3 - i(\omega_2^1 e_1 + \omega_2^3 e_3) \;=\; -i\omega_2^1(e_1 - ie_2) + (\omega_1^3 - i\omega_2^3)e_3, \\ &=\; (u_z\, dz - u_{\bar{z}}\, d\bar{z})(e_1 - ie_2) + (h\, e^u dz + He^u d\bar{z})e_3 \end{aligned}$$

$$\begin{aligned} \Longrightarrow \qquad d\mathbf{x}_z \;&=\; \frac{1}{2}\, d(e^u)(e_1 - ie_2) + \frac{1}{2}\, e^u d(e_1 - ie_2) \\ &=\; \frac{1}{2}\, e^u(u_z\, dz + u_{\bar{z}}\, d\bar{z})(e_1 - ie_2) + \frac{1}{2}\, e^u(u_z\, dz - u_{\bar{z}}\, d\bar{z})(e_1 - ie_2) \\ &\qquad\quad + (h\, e^u dz + He^u d\bar{z})e_3 \\ &=\; e^u(u_z\, dz)(e_1 - ie_2) + \frac{1}{2}\, e^u(h\, e^u dz + He^u d\bar{z})e_3 \end{aligned}$$

Since $d\mathbf{x}_z = \mathbf{x}_{zz}dz + \mathbf{x}_{z\bar{z}}d\bar{z}$, we collect the coefficients of $d\bar{z}$ above to conclude that

$$\mathbf{x}_{z\bar{z}} = \frac{1}{2}\, e^{2u} He_3$$

so that $\mathbf{x}_{z\bar{z}} = 0$ if and only if $H \equiv 0$, which by the Lagrange-Meusnier Theorem is equivalent to the minimality of $\mathbf{x}$. □

The key observation for what remains is that $\mathbf{x} : M^2 \to \mathbf{R}^3$ is harmonic if and only if $\mathbf{x}_z : M^2 \to \mathbf{C}^3$ holomorphic. Since $z$ is the complex structure induced by $\mathbf{x}$, the bidegree decomposition of the metric shows that

$$e^{2u} dz d\bar{z} \;=\; ds^2 \;=\; d\mathbf{x}\cdot d\mathbf{x} \;=\; \mathbf{x}_z\!\cdot\!\mathbf{x}_z\, dzdz \;+\; 2\,\mathbf{x}_z\!\cdot\!\mathbf{x}_{\bar{z}}\, dzd\bar{z} \;+\; \mathbf{x}_{\bar{z}}\!\cdot\!\mathbf{x}_{\bar{z}}\, d\bar{z}d\bar{z}$$

$$\Longrightarrow \quad \begin{cases} \mathbf{x}_z\!\cdot\!\mathbf{x}_z \;=\; \mathbf{x}_{\bar{z}}\!\cdot\!\mathbf{x}_{\bar{z}} \;=\; 0 \\ \mathbf{x}_z\!\cdot\!\mathbf{x}_{\bar{z}} \;=\; \tfrac{1}{2}e^{2u} \end{cases}$$

Thus $\mathbf{x}$ is minimal when $\mathbf{x}_z$ is a *null* (or *isotropic*[10]) holomorphic map $\mathbf{x}_z\!\cdot\!\mathbf{x}_z \equiv 0$, or equivalently, when the the differential $d\mathbf{x}$ has a non-vanishing, null, holomorphic (1,0)-bidegree $\boldsymbol{\alpha} := \mathbf{x}_z dz \in \bigwedge^{1,0}(M^2; \mathbf{C}^3)$, with $\boldsymbol{\alpha}\cdot\boldsymbol{\alpha} \equiv 0$. For brevity, a non-vanishing, null $\mathbf{C}^3$-valued holomorphic 1-form on $M^2$ is called an *abelian differential.*

Weierstrass was able to parameterize the space of abelian differentials [23]. Each can be represented as a product of a null meromorphic $\mathbf{C}^3$-valued map and a holomorphic 1-form: If $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ with $\alpha_i$ holomorphic 1-forms, the dot product can be written $\boldsymbol{\alpha}\cdot\boldsymbol{\alpha} = \alpha_1^2 + \alpha_2^2 + \alpha_3^2 = (\alpha_1 + i\alpha_2)(\alpha_1 - i\alpha_2) + \alpha_3^2$, so that nullity of $\boldsymbol{\alpha}$ becomes the condition that $(\alpha_1 + i\alpha_2)(\alpha_1 - i\alpha_2) = -\alpha_3^2$. Using this relation we can express the entries $\alpha_i$ in terms of a meromorphic function $g : M^2 \to \mathbf{C} \cup \{\infty\}$ and holomorphic 1-form $\eta \in \bigwedge^{1,0}(M^2; \mathbf{C})$ defined by

$$\alpha_1 - i\alpha_2 \;=:\; \eta \qquad \text{and} \qquad \frac{\alpha_3}{\alpha_1 - i\alpha_2} \;=:\; g, \qquad \text{so that}$$

$$\alpha_1 + i\alpha_2 \;=\; \frac{-\alpha_3^2}{\alpha_1 - i\alpha_2} \;=\; -g^2\eta$$

$$\Longrightarrow \quad \begin{cases} 2\alpha_1 \;=\; \eta - g^2\eta \\ 2\alpha_2 \;=\; i(\eta + g^2\eta) \\ \;\alpha_3 \;=\; g\eta \end{cases} \qquad \Longrightarrow \qquad \boldsymbol{\alpha} \;=\; \begin{pmatrix} \frac{1}{2}(1-g^2) \\ \frac{i}{2}(1-g^2) \\ g \end{pmatrix} \eta$$

In order for $\boldsymbol{\alpha}$ to be non-vanishing holomorphic, we see that the poles and zeros of $g$ and $\eta$ must satisfy a "balancing condition:"

[10] Given a vector space $V$ over a field $k$ equipped with a quadratic form $Q : V \to k$, we say that a vector $\mathbf{v} \in V$ is *null*, or *isotropic*, if $Q(v) = 0$. In the present context, $V = \mathbf{C}^3$ and $Q$ is the quadratic form associated to the dot product, $Q(\mathbf{v}) := \mathbf{v}\cdot\mathbf{v}$.

$$\eta \text{ has a zero of order } 2m \text{ at } p \quad \Leftrightarrow \quad g \text{ has a pole of order } m \text{ at } p \qquad \text{(B)}$$

and provided that the zeros of $\eta$ are discrete,[11] any $\boldsymbol{\alpha}$ can be represented in this way. Conversely, any meromorphic function and holomorphic 1-form pair $(g, \eta)$ satisfying the balancing condition (B) yields an abelian differential via the formula above. This representation theorem really parameterizes all minimal immersions of a given surface $M^2$, since abelian differentials are exactly the (1,0)-parts of differentials of minimal immersions:

**Theorem 12.** *(Enneper-Weierstrass) Given a minimal immersion $\mathbf{x} : M^2 \to \mathbf{R}^3$, there exists a meromorphic function and holomorphic 1-form pair $(g, \eta)$ on $M^2$ satisfying condition* (B) *such that $\mathbf{x}$ is given by real part of the antiderivative of $\boldsymbol{\alpha}$:*

$$\mathbf{x}(z) \;=\; Re \int \mathbf{x}_z dz \;=\; Re \int \boldsymbol{\alpha} \;=\; Re \int_{z_0}^{z} \begin{pmatrix} \frac{1}{2}(1-g^2) \\ \frac{i}{2}(1-g^2) \\ g \end{pmatrix} \eta$$

*Conversely, a meromorphic function and holomorphic 1-form pair $(g, \eta)$ on $M^2$ satisfying the balancing condition* (B) *induce a minimal immersion $\mathbf{x} = Re \int \boldsymbol{\alpha} : \widetilde{M}^2 \to \mathbf{R}^3$ of the universal cover $\pi : \widetilde{M}^2 \to M^2$. This immersion descends to $M^2$ if the "real periods" $Re \int_\gamma \boldsymbol{\alpha}$ vanish for all generators $\gamma$ of the fundamental group $\pi_1(M^2)$.*

This is the famous *Weierstrass representation* for minimal surfaces, and the pair $(g, \eta)$ are called *Weierstrass data* for the immersion constructed from the recipe in the theorem. The antiderivative $\int_\gamma \boldsymbol{\alpha}$ is evaluated as a line integral along any curve $\gamma$ joining the fixed base point $z_0$ to $z$, and the condition that $Re \int_\gamma \boldsymbol{\alpha} = 0$ for all $\gamma \in \pi_1(M^2)$ is simply that $Re \int \boldsymbol{\alpha}$ is path independent. This theorem replaces a

[11] In the case $\alpha_1 - i\alpha_2 \equiv 0$ identically on $M^2$, the representation statement must be rephrased: there exists a 1-form $\eta$ such that $\boldsymbol{\alpha} = (1, i, 0)^T \eta$.

second order PDE (the MSE) with a pair of first order PDEs (the Cauchy-Riemann equations), vastly simplifying the task of generating new minimal surfaces. Basic examples tend to have pleasantly basic Weierstrass data $(g, \eta)$: a plane has data set $(g_0, \eta)$ on $M^2 = \mathbf{C}$ with $g_0 \in \mathbf{C}$ constant and $\eta$ arbitrary; the Catenoid and Helicoid are given by $(\frac{1}{z}, dz)$ and $(\frac{i}{z}, dz)$, respectively, on $M^2 = \mathbf{C} - \{0\}$. In general, one can attempt to build minimal immersions of $M^2$ by specifying a set of Weierstrass data and then solving the "period problems," that is, adjusting the data in such a way that the real periods become zero. When successful, this technique can provide surprising control over several characteristics of the resulting immersion, since all of fundamental geometric quantities associated to the surface can be expressed in terms of the data directly. Recall that the 2-sphere $S^2 \subset \mathbf{R}^3$ can be identified with the extended complex plane $\mathbf{C} \cup \{\infty\}$ via stereographic projection $\sigma : S^2 \to \mathbf{C} \cup \{\infty\}$, $\sigma(u, v, w) = \frac{u+iv}{1-w}$:

**Proposition 13.** *Let $\mathbf{x} : M^2 \to \mathbf{R}^3$ is a minimal immersion with Weierstrass data $(g, \eta)$ and Gauss map $e_3 : M^2 \to S^2$. Then $g$ is the composition $g = \sigma \circ e_3$, and the fundamental forms $I = ds^2$ and $II = 2Re(II^{2,0})$ are given by $ds^2 = (1 + |g|^2)^2 \, |\eta|^2$ and $II^{2,0} = -\eta \, dg$.*

We will rely heavily on a variant of the Weierstrass representation called the *spinor representation*. For our purposes a completely local description will suffice. Note that the *spinor map* $\varphi : \mathbf{C}^2 \to \mathbf{C}^3$ defined by

$$\varphi \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(q^2 - p^2) \\ \frac{i}{2}(q^2 + p^2) \\ qp \end{pmatrix}$$

is null-valued, and if $p, q : M^2 \to \mathbf{C}$ are holomorphic functions with no common zeros, then locally the combination $\varphi\binom{p(z)}{q(z)}dz$ is a well-defined abelian differential. Since the holomorphic 1-form $\eta$ can locally be written $\eta = f(z)dz$ for some holomorphic function $f$, we can always locally express an abelian differential in this form, where $g = \frac{p}{q}$ and $f = q^2$. The functions $(p, q)$ are called *spinor fields* or *spinor data*[12] for the surface $\mathbf{x} = Re \int \varphi \binom{p(z)}{q(z)} dz$. While perhaps lacking the geometric significance of the data $(g, \eta)$, the spinor representation puts the two data components on a more equal footing, simplifies certain computations and, as well shall see, provides the mechanism by which an analogy between surfaces and particles emerges.

We conclude by mentioning an extremely important subclass of minimal surfaces which was studied extensively by Robert Osserman [16]: the *complete* minimal surfaces. These are interesting for many reasons, but the main result, due to Osserman, is that they must have *finite topology*, that is, $M^2$ is conformal to a compact surface $S$ with finitely many punctures, $M^2 \simeq S - \{p_1, ..., p_n\}$. Now, the total curvature of $M^2$ is simply the negative area of the image $g(M^2)$ in $S^2$ (computed with multiplicity of course). If this number is finite, the Gauss map must extend across the punctures to a well-defined meromorphic function on $S$ compact, $g : S \to S^2$. Then the degree of $g$ is a well-defined integer $deg(g)$, and the total curvature is $-4\pi\, deg(g)$. Thus the complete minimals form a natural class of non-compact[13] surfaces whose total curvature $\iint K\, dA$ is *quantized* (that is, the total curvature only assumes a discrete set of values), and we can think of this quantization as an analogue of the Gauss-Bonnet Theorem for a special non-compact case.

[12]There is a more precise global use of the term *spinor field* for surfaces, which has played an important role in much of the current research in surface theory, but this technical refinement is too involved for our purposes.

[13]By Liouville's theorem, a minimal surface $M^2$ cannot be compact, since otherwise the holomorphic function $\mathbf{x}_z$ would be constant.

## 1.4 CMC1 surfaces: the Bryant representation

Look back at the complex structure equations for immersions into the space forms $S^3_\epsilon$:

$$\begin{cases} -4e^{-2u}u_{z\bar{z}} \;=\; H^2 - |\,h|^2 + \epsilon \\ \;h_{\bar{z}} + 2h\,u_{\bar{z}} \;=\; H_z \end{cases}$$

Note that the Codazzi equation is already independent of $\epsilon$, and if the mean curvature is constant, $H \equiv c$, it becomes $h_{\bar{z}} + 2h\,u_{\bar{z}} \;=\; 0$, which is independent of $H$ as well. If a CMC-H immersion $x : M^2 \to S^3_\epsilon$ has invariants $\{H, u, h\}$, then $\{\widetilde{H}, u, h\}$ will be a set of invariants for a CMC-$\widetilde{H}$ immersion $\tilde{x} : M^2 \to S^3_{\tilde{\epsilon}}$ with $\tilde{\epsilon} = H^2 + \epsilon - \widetilde{H}^2$, where both sets of invariants satisfy the *same* structure equations. Since these equations uniquely determine the surface up of rigid motion, this means that there is a local correspondence between CMC-H surfaces in $S^3_\epsilon$ and CMC-$\widetilde{H}$ surfaces in $S^3_{\tilde{\epsilon}}$ when $H^2 + \epsilon = \widetilde{H}^2 + \tilde{\epsilon}$. This is called the *Lawson correspondence.*[14]

The special case of interest to us is $H^2 + \epsilon = 0$, and for reasons we will soon explain, CMC-H immersions into $S^3_{\tilde{\epsilon}}$ satisfying this condition are called *horospherical surfaces.* In particular, when $\epsilon = 0$ (Euclidean space) we have $H \equiv 0$ (minimal), and when $\tilde{\epsilon} = -1$ (hyperbolic space) then $\widetilde{H} \equiv 1$, so that we have the local correspondence

$$\big\{\mathbf{x} : M^2 \to \mathbf{R}^3 \quad \text{minimal}\big\} \quad \longleftrightarrow \quad \big\{f : M^2 \to H^3 \quad \text{CMC1}\big\}$$

For brevity, we will frequently call a surface simply "CMC1" when it is a CMC1 surface in hyperbolic space, and a pair of corresponding minimal and CMC1 surfaces will be called *cousins.* Horospheres[15] are the most trivial examples of CMC1 surfaces,

[14] The terminology is due to Lawson's use of this phenomena in his work on surfaces in $S^3$, but this is not entirely justified historically, since the correspondence was well-known to Bianchi and Darboux.

[15] A *horosphere* is a (metric) sphere in $H^3$ which is tangent to the infinity boundary $\partial_\infty H^3$, or equivalently, a sphere whose center is a point on the infinity boundary. Note that these are not spheres topologically since the point of tangency, or "point at infinity," is missing.

and can be thought of as CMC1 analogues of Euclidean planes. That horospheres are naturally occurring objects of hyperbolic geometry already suggests that CMC1 surfaces should be a natural category of surfaces in their own right, but their correspondence to a class as important as minimal surfaces singles them out as potentially very interesting.

Such were probably the sentiments of Robert Bryant, who predicted that CMC1 surfaces might share a great deal of the special geometric behavior exhibited by their minimal cousins. He began his study [4] by first showing that, like minimal surfaces, they possess a kind of "Weierstrass representation" in terms of holomorphic data. He relied on the *Hermitian matrix model* of hyperbolic space, the existence of which is a special feature of dimension three: Consider the vector space of $2\times 2$ Hermitian matrices $Herm_2(\mathbf{C}) := \{X \in M_2(\mathbf{C}) \mid X = X^*\}$, and define the quadratic form $Q(X) = -det(X)$, which polarizes to a bilinear form $\langle\cdot,\cdot\rangle$. The Pauli spin matrices $\sigma_i$ form an orthonormal basis for $Herm_2(\mathbf{C})$:

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

That $(Herm_2(\mathbf{C}), \langle\cdot,\cdot\rangle)$ is isometric to Minkowski space $\mathbf{R}^{1,3}$ follows by checking that the "Pauli map"

$$\begin{array}{rccc} S: & \mathbf{R}^{1,3} & \longrightarrow & Herm_2(\mathbf{C}) \\ & x^i e_i & \longmapsto & x^i \sigma_i \end{array}$$

is in fact an isometry. Then $\langle\cdot,\cdot\rangle$ is a pseudo-metric on $Herm_2(\mathbf{C})$ of signature (1,3), and $H^3_{Herm} := \{X \in M_2(\mathbf{C}) \mid X = X^*,\ detX = 1,\ trX > 0\}$ is (one sheet of) the hyperboloid of radius $-1$. So $\langle\cdot,\cdot\rangle$ restricts to a metric of constant curvature $-1$, and $H^3_{Herm}$ is indeed a model for hyperbolic 3-space $H^3$. The reason this model is desirable is because of the particular representation of the isometry group: $SL_2(\mathbf{C})$

acts transitively and isometrically on $H^3_{Herm}$ by conjugation, $A \cdot X = AXA^*$, with isotropy subgroup $SU_2$ at $\sigma_0 \in H^3_{Herm}$. Thus $H^3 \simeq SL_2(\mathbf{C})/SU_2$ and we let $\pi : SL_2(\mathbf{C}) \to H^3$ be the natural projection $\pi(A) = AA^*$.

If $f : M^2 \to H^3$ is an immersion of a surface, a *frame* along $f$ is just a lift, $F : M^2 \to SL_2(\mathbf{C})$, with $f = FF^*$. Bryant observed that if $f$ is CMC1 then it has a frame $F$ which is holomorphic and *null*, in the sense that $det(F^{-1}dF) = 0$, or equivalently, $\tau = F^{-1}dF$ is a non-vanishing, null $\mathfrak{sl}_2(\mathbf{C})$-valued holomorphic 1-form on $M^2$. As before, 1-forms $\tau$ of this type can be factored as a null $\mathfrak{sl}_2(\mathbf{C})$-valued holomorphic map times an arbitrary 1-form:

$$\tau = \begin{pmatrix} g & -g^2 \\ 1 & -g \end{pmatrix} \eta$$

where $(g, \eta)$ must satisfy the balancing condition (B) from section 3 and are possibly multi-valued. The assertion above, together with its converse, constitute the *Bryant representation*, which we state here without proof:

**Theorem 14.** *(Bryant) Given a CMC1 immersion $f : M^2 \to H^3$, there exists data $(g, \eta)$ on $M^2$ and a null holomorphic frame $F : M^2 \to SL_2(\mathbf{C})$ satisfying $det(F^{-1}dF) = \tau$, with $\tau$ defined as above. Conversely, given data $(g, \eta)$ on $M^2$, if $F : M^2 \to SL_2(\mathbf{C})$ is a holomorphic solution of the ODE system $F^{-1}dF = \tau$, then the projection $f = FF^*$ is CMC1.*

For conceptual unity, and with all due respect to Bryant, we may sometimes refer to this as the Weierstrass representation also, and again call $(g, \eta)$ Weierstrass data for the CMC1 immersion $f$ constructed in the theorem. A computation shows that this immersion has fundamental forms $I = ds^2$ and $II = 2Re(II^{2,0}) + ds^2$ given by $ds^2 = (1 + |g|^2)^2\, |\eta|^2$ and $II^{2,0} = -\eta\, dg$. Thus minimal and CMC1 surfaces with the same data correspond. In both cases the Weierstrass representation only determines a surface up to some ambiguity: a minimal surface is determined by $(g, \eta)$ up to translation, and a CMC1 surface is determined up to a hyperbolic rigid motion. This is indeed consistent with the Lawson correspondence, which is only an existence result and does not distinguish among congruent surfaces.

It is important to note that the Gauss map of the CMC1 space with data $(g, \eta)$ is not simply $g$. In fact, it is not even clear how the Gauss map for a surface in hyperbolic space should be defined, which in Euclidean space $\mathbf{R}^3$ relies on parallel translation to consistently identify the unit vectors $e_3$ at different points on the surface with points in $S^2$. Instead, we can use parallel translation in $H^3$ to identify unit vectors at different points on the surface with points in the infinity boundary $\partial H^3$: let $\mathbf{n}$ be the unit normal at a point $p \in M^2$, and consider the (unique) geodesic $\gamma(s)$ such that $\gamma(0) = p$ and $\gamma'(0) = \mathbf{n}$. This curve terminates on the infinity boundary at the point $g_h(p) := \lim_{s\to\infty} \gamma(s)$, and if $\partial H^3$ is regarded as the space of directions in $H^3$, then $g_h : M^2 \to \partial H^3$ is indeed the analogue of the Gauss map, called the *hyperbolic Gauss map*.

In the language of Möbius geometry, the hyperbolic Gauss map coincides with the *mean curvature sphere congruence*, the map that assigns to each point the unique tangent sphere with the same mean curvature as the surface at that point. In the CMC1 case this sphere must be a horosphere, and the hyperbolic Gauss map simply gives the center of this sphere. Regarding planes in $\mathbf{R}^3$ as zero mean curvature (horo)spheres, the Gauss map of a minimal surface can also be identified with this sphere congruence. In fact, this is a general Möbius-geometric characterization of CMC-H immersions $f : M^2 \to S^3_\epsilon$ with $H^2 + \epsilon = 0$, which inspired Hertrich-Jeromin [11] to call these sets of surfaces collectively the *horospherical surfaces*.

In the hyperboloid model, the infinity boundary can be identified with the light cone (or rather, the projectivized light cone), which is the set of null vectors in

Minkowski space. In the Hermitian matrix model, null vectors are matrices with zero determinant, which when projectivized is

$$\partial_\infty H^3_{Herm} \;\simeq\; \{\,[X] \mid \; detX = 0,\;\; X \sim \lambda X \;\; \forall\, \lambda \in \mathbf{R}^*\}$$

Given an immersion $f = FF^* : M^2 \to H^3$ with frame $F = (F_1, F_2) : M^2 \to SL_2(\mathbf{C})$, where $F_1, F_2 : M^2 \to \mathbf{C}^2$ are the columns of $F$, the hyperbolic Gauss map can be expressed $g_h = [\dot{F}_1^* \dot{F}_1]$. From this definition one can check that if $g$ is constant, then $g_h$ is constant. Since $g_h$ is only constant on horospheres, these are indeed the cousins of planes. Pursuing the minimal/CMC1 analogy further, Bryant proved

**Proposition 15.** *Assuming $f : M^2 \to H^3$ is not totally umbilic, then $f$ is CMC1 if and only if $g_f$ is conformal.*

The Bryant representation helps facilitate calculations needed to obtain results like this one, and greatly assists in understanding the minimal/CMC1 correspondence. But there still seems to be a general lack of computable examples of minimal/CMC1 cousins. The central difficulty is that the Weierstrass data, which merely need to be integrated to parameterize a given minimal surface, appear as *coefficients* in an ODE that needs to be solved to parameterize the CMC1 cousin. We will attempt to generate new examples by two methods: first, by searching for examples whose data gives rise to well-studied ODE, and second, by considering the transformations of these surfaces, which we now explain.

## 1.5 Transformation theory

Consider again a "Plateau soap bubble" $M^2$, whose boundary is a loop of wire $C$. If the wire is made to bend slowly into a new shape $\widetilde{C}$, the bubble $M^2$ will also bend so as to remain approximately minimal and still contain the wire as its boundary, resulting in a new surface $\widetilde{M}^2$, which might be quite different from the original. The study of how classes of submanifolds can (or cannot) be perturbed so as to preserve (or modify) their geometry is a general area of study which we call *transformation theory*. A *transformation* of a class of submanifolds $\mathcal{C}$ is just a map $T : \mathcal{C} \to \mathcal{C}$, typically continuous with respect to some topology[16] on $\mathcal{C}$, though not necessarily invertible. We denote the set of transformations on $\mathcal{C}$ by $Trans(\mathcal{C})$, which can be regarded as a monoid with respect to composition. A *deformation*[17] is a curve $\rho : \mathbf{R} \to Trans(\mathcal{C})$ such that $\rho(0) = Id$, also usually continuous in some sense. In this way, $x_t := \rho(t)x$ will be a 1-parameter family of manifolds which *deform* an initial manifold $x = x_0 \in \mathcal{C}$.

Homothety is an $\mathbf{R}^+$-deformation of submanifolds $N \subset \mathbf{R}^n$ in Euclidean space, $\rho_t(N) := t(N)$. A less trivial example for CMC surfaces uses Bonnet's theorem: suppose a CMC-$H$ surface $\mathbf{x}$ has fundamental forms $I$ and $II = 2\,Re(II^{2,0}) + HI$, and let $\mathbf{x}_\theta$ be the *associate* surface defined by forms $I$ and $II_\theta = 2\,Re(e^{i\theta}II^{2,0}) + HI$. Then $\rho_\theta(\mathbf{x}) := \mathbf{x}_\theta$ is an isometric $S^1$-deformation. In the case of minimal surfaces, this deformation can be equivalently defined be letting $\mathbf{x}_\theta$ be the surface determined by Weierstrass data $(g, e^{i\theta}\eta)$ where $(g,\eta)$ is the data for $\mathbf{x}$. Note that $\mathbf{x}_{\pi/2}$ is the conjugate surface,[18] and $\mathbf{x}_\theta$ is a periodic 1-parameter family of surfaces, all of which are isometric.

*Example* 1. Inspecting the Weierstrass data of the Catenoid and Helicoid, we see that they are indeed conjugate, so the associates deform one into the other:

[16] There does not seem to be a universally preferred topology on an arbitrary set of manifolds or submanifolds. Usually, the nature of the objects themselves suggest a natural topology, and in the case of minimal surfaces there is a variety of options, e.g. [17].

[17] More generally, an action of a Lie group $G$ on $\mathcal{C}$ is a *$G$-deformation*, i.e. a homomorphism $\rho : G \to Trans(\mathcal{C})$, also called an *$n$-parameter deformation* when $dim\,G = n$.

[18] The *conjugate* of a minimal surface $\mathbf{x}$ is that minimal surface $\mathbf{y}$ such that $\mathbf{x} + i\mathbf{y}$ is holomorphic, that is, the coordinate functions of $\mathbf{y}$ are the harmonic conjugates of those of $\mathbf{x}$

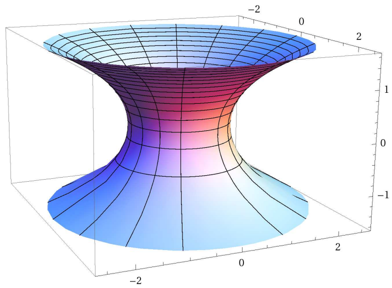

**Fig.1.3** $\theta = 0$

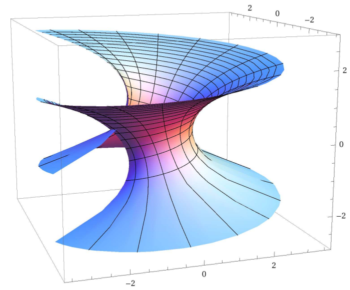

**Fig.1.4** $\theta = \frac{\pi}{10}$

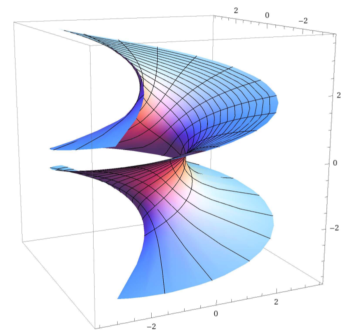

**Fig.1.5** $\theta = \frac{\pi}{4}$

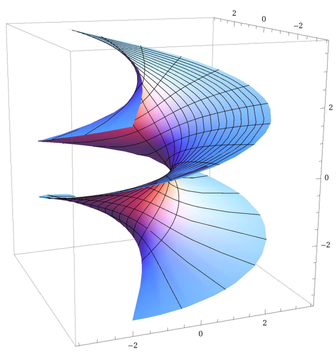

**Fig.1.6** $\theta = \frac{4\pi}{10}$

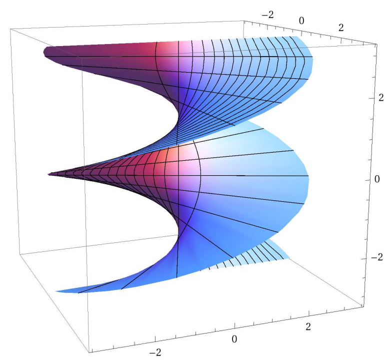

**Fig.1.7** $\theta = \frac{\pi}{2}$

More generally, let a minimal surface $\mathbf{x}$ have data $(g, \eta)$ and let $f(z)$ be a non-vanishing holomorphic function on $M^2$. If $\mathcal{A}$ denotes the space of minimal surfaces and $\mathbf{x}_f$ denotes the surface determined by data $(g, f(z)\eta)$, we call the map

$$\begin{aligned} T_f \,:\, \mathcal{A} &\rightarrow \mathcal{A} \\ \mathbf{x} &\mapsto \mathbf{x}_f \end{aligned}$$

the *conformal transform* associated to $f(z)$. By Prop 12 this transform leaves the Gauss map fixed and changes the metric by conformal factor $|f(z)|^2$, so $\mathbf{x}$ and $\mathbf{x}_f$ are conformally related. The surfaces $\mathbf{x}_f$ are called the *conformal associates*, the associate surfaces being a special case.

The extent to which horospherical surfaces can be deformed is perhaps the main concern of this work. Deforming the boundary of a compact Plateau bubble bends the interior surface, but our minimal surfaces have no boundary: Is there an analogous deformation for these non-compact surfaces? Letting $\mathcal{B}$ denote the set of CMC1 surfaces, Lawson's correspondence provides a bijection (that is, as moduli spaces) $\mathcal{L} : \mathcal{A} \to \mathcal{B}$ and one might next ask: Given a deformation $\rho$ of $\mathcal{A}$, what is the geometry of the deformation $\mathcal{L} \circ \rho \circ \mathcal{L}^{-1}$ on $\mathcal{B}$? That is, how does the transformation theory on $\mathcal{A}$ carry over to $\mathcal{B}$? Some experience with generating examples quickly shows that the question is intractable when stated in this kind of generality. For example, the conformal transform on the CMC1 side is defined by $T_f(x) = x_f$, where $f(z)$ is holomorphic on $M^2$ and $x_f$ is the CMC1 with data $(g, f(z)\eta)$ when $x$ is the CMC1 with data $(g, \eta)$. By Bryant's theorem, this transform is *equivariant*, in the sense that $\mathcal{L} \circ T_f = T_f \circ \mathcal{L}$. Yet computationally, its behavior is wildly different on CMC1 surfaces than on minimal surfaces, since the ODE $F^{-1}dF = \tau$ may have little in common with the ODE $\tilde{F}^{-1}d\tilde{F} = f(z)\tau$, even when $f$ is constant. In the next chapter we will study an equivariant deformation for which this question *can* be approached.

# Chapter 2

# The Goursat Transform for Minimal Surfaces

## 2.1 The group of complex rotations

It will streamline our theory to refer to minimal surfaces via their associated *minimal curves*, defined as follows.

**Definition 16.** A *minimal curve* is a regular holomorphic map $\boldsymbol{\gamma} : M^2 \to \mathbf{C}^3$ such that the velocity $\dot{\boldsymbol{\gamma}}(p)$ is null at each point $p \in M^2$ (i.e. if $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \gamma_3)$, then $\dot{\boldsymbol{\gamma}} \cdot \dot{\boldsymbol{\gamma}} = \sum_{i=1}^{3} \dot{\gamma}_i^2 \equiv 0$).

Another way to say this is that a minimal curve is a holomorphic map from $M^2$ to $\mathbf{C}^3$ whose derivative takes values in the affine quadric $Q_1$ of non-zero null vectors

$$Q_1 := \big\{(z_1, z_2, z_3) \in \mathbf{C}^3 - \{\mathbf{0}\} \mid z_1^2 + z_2^2 + z_3^2 = 0\big\}.$$

Note that the real part $\mathbf{x} : M^2 \to \mathbf{R}^3$ of a minimal curve $\boldsymbol{\gamma} = \mathbf{x} + i\mathbf{y} : M^2 \to \mathbf{C}^3$ is a minimal immersion whose conjugate is the imaginary part $y : M^2 \to \mathbf{R}^3$ of $\boldsymbol{\gamma}$, hence the name. Conversely, given any minimal immersion $\mathbf{x} : M^2 \to \mathbf{R}^3$ with conjugate $\mathbf{y} : M^2 \to \mathbf{R}^3$, the combination $\boldsymbol{\gamma} := \mathbf{x} + i\mathbf{y}$ defines a minimal curve (the minimal curve *associated* to $\mathbf{x}$). Thinking of $Re : \mathbf{C}^3 \to \mathbf{R}^3$ as a projection operator, we may regard minimal surfaces as projections of minimal curves.

Consider the complexification of the Euclidean rotation group $SO_3(\mathbf{R})$,

$$SO_3(\mathbf{C}) := \big\{O \in GL_3(\mathbf{C}) \mid O^T O = I\big\}.$$

Since these "complex rotations" preserve the complexified Euclidean inner product on $\mathbf{C}^3$, they also preserves the quadric $Q_1$. Therefore, the full complexified Euclidean group of motions $E_3(\mathbf{C}) := SO_3(\mathbf{C}) \ltimes \mathbf{C}^3$ acts on minimal curves according to $(O, \mathbf{v}) \cdot \boldsymbol{\gamma} := O\boldsymbol{\gamma} + \mathbf{v}$, analogous to the way the Euclidean group $E_3(\mathbf{R})$ acts on 1-dimensional curves $\boldsymbol{\gamma} : \mathbf{R} \to \mathbf{R}^3$. This led E. Goursat [10] to study the following transformation of minimal surfaces:

**Definition 17.** Let $\mathbf{x}\colon M^2 \to \mathbf{R}^3$ be a minimal surface with associated minimal curve $\boldsymbol{\gamma} : M^2 \to \mathbf{C}^3$. A *Goursat transform* of $\mathbf{x}$ is a minimal surface $\mathbf{x}_{(O,\mathbf{v})} := Re\,(O\boldsymbol{\gamma} + \mathbf{v})$.

Since the Weierstrass representation only defines a minimal surface up to an additive constant, we will often only consider the $SO_3(\mathbf{C})$ factor of $E_3(\mathbf{C})$, denoting the transform simply by $\mathbf{x}_O$. This provides $\mathbf{R}$-deformations $\mathbf{x}_t$ of a minimal surface $\mathbf{x}$ by defining $\mathbf{x}_t := \mathbf{x}_{a(t)}$ where $a : \mathbf{R} \to SO_3(\mathbf{C})$ is any smooth map such that $a(0) = I$. Since $SO_3(\mathbf{C})$ has real dimension 6, $SO_3(\mathbf{R})$ occupying 3 of these, we regard the Goursat transform as providing a 3-parameter family of deformations. Note also, in the same way that a constant real conformal transform is homothety, the Goursat transform restricted to $E_3(\mathbf{R}) \subset E_3(\mathbf{C})$ is simply rigid motion. Similarly, as a constant *complex* conformal transform by $f(z) = c = a + ib$ produces an arbitrary linear combination of an immersion and its conjugate, $\mathbf{x} \mapsto \mathbf{x}_c = a\mathbf{x} - b\mathbf{y}$, the Goursat transform provides a kind of combination of these surfaces, since if $O \in SO_3(\mathbf{C})$ has real and imaginary parts $O = A + iB$, then $\mathbf{x} \mapsto \mathbf{x}_O = A\mathbf{x} - B\mathbf{y}$.

One might try to study the real and imaginary parts of $SO_3(\mathbf{C})$ matrices in an effort to understand this transformation, but Goursat decided instead to pursue a decomposition theorem [10] for $SO_3(\mathbf{C})$. The result is a clever factorization of a matrix $O \in SO_3(\mathbf{C}) - SO_3(\mathbf{R})$ into rotations by real and "imaginary" angles: Goursat proved that there is a unique (up to sign) real unit vector $\mathbf{v} \in \mathbf{R}^3$ such that $O\mathbf{v} \in \mathbf{R}^3$ is also real. If we complete $\mathbf{v}$ to a $\mathbf{C}$-basis for $\mathbf{C}^3$ and choose $R \in SO_3(\mathbf{R})$ such that $RO\mathbf{v} = \mathbf{v}$, then with respect to this new basis the matrix representation of $RO$ has the form:

$$RO = \begin{pmatrix} \cos z & -\sin z & 0 \\ \sin z & \cos z & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos x & -\sin x & 0 \\ \sin x & \cos x & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos iy & -\sin iy & 0 \\ \sin iy & \cos iy & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

for some complex number $z = x + iy$. The first factor is just a rotation about the $\mathbf{v}$-direction by an angle $x$, while the second is what one would mean by "rotation by an *imaginary* angle $iy$," or what is also called "hyperbolic rotation" since $\cos iy = \cosh y$, $\sin iy = -i \sinh y$.

Perhaps more fascinating than Goursat's decomposition theorem is the method used to prove it. One way to prove the existence of the real unit vector $\mathbf{v}$ involves exhibiting $SL_2(\mathbf{C})$ as the double (universal) cover of $SO_3(\mathbf{C})$, which we now describe following [12]. Recall the extended stereographic projection $\sigma : S^2 \to \mathbf{C}P^1$ discussed in the introduction. This gives rise to the double cover of $SO_3(\mathbf{R})$ by $SU_2$

$$\begin{aligned} \Sigma \ : \ SU_2 &\longrightarrow SO_3(\mathbf{R}) \\ A \ &\longmapsto \sigma^{-1} \circ A \circ \sigma \end{aligned}$$

Inspired by this picture, we seek an extension of $\sigma$ to define a double covering of $SO_3(\mathbf{C})$ by $SL_2(\mathbf{C})$. First we need appropriate objects for these groups to act on. While $SL_2(\mathbf{C})$ still acts (non-isometrically) on $\mathbf{C}P^1$ by projective transformation, $SO_3(\mathbf{C})$ does not seem to act on $S^2 = \{\mathbf{x} = (x_1, x_2, x_3) \in \mathbf{R}^3 \mid x_1^2 + x_2^2 + x_3^2 = 1\}$ at all. It *does* however act on a kind of "complexified" 2-sphere, the affine quadric

$$S^2_{\mathbf{C}} := \{\mathbf{z} = (z_1, z_2, z_3) \in \mathbf{C}^3 \mid z_1^2 + z_2^2 + z_3^2 = 1\}.$$

Now we need something larger to replace $\mathbf{C}P^1$, since $S^2_{\mathbf{C}}$ is clearly too big to be covered by $\mathbf{C}P^1$ alone. In fact, we almost need *two* $\mathbf{C}P^1s$, or rather, the space of point pairs,

$$\mathbf{C}P^1_{\mathcal{P}} := \big\{(u,v) \in \mathbf{C}P^1 \times \mathbf{C}P^1 \mid u \neq v\big\},$$

on which $SL_2(\mathbf{C})$ acts diagonally. Note that the 2-sphere and projective line are natural subsets $S^2 \subset S^2_{\mathbf{C}}$ and $\mathbf{C}P^1 \cong \big\{(u,v) \mid v = -1/\bar{u}\big\} \subset \mathbf{C}P^1_{\mathcal{P}}$. The desired identification of these objects is then

$$\begin{aligned} \tilde{\sigma} \,:\, S^2_{\mathbf{C}} &\longrightarrow \mathbf{C}P^1_{\mathcal{P}} \\ \mathbf{z} &\longmapsto (\tilde{\sigma}_N(\mathbf{z}), -\tilde{\sigma}_S(\mathbf{z})) \end{aligned}$$

where $\tilde{\sigma}_N$ and $\tilde{\sigma}_S$ are north and south "complex" stereographic projections, defined by $\tilde{\sigma}_N(\mathbf{z}) = \begin{bmatrix} z_1 + iz_2 \\ 1 - z_3 \end{bmatrix}$, $\tilde{\sigma}_S(\mathbf{z}) = \begin{bmatrix} z_1 + iz_2 \\ 1 + z_3 \end{bmatrix}$. Indeed, $\tilde{\sigma}$ is a biholomorphic equivalence, with inverse $\tilde{\sigma}^{-1}\big(\begin{bmatrix} u \\ 1 \end{bmatrix}, \begin{bmatrix} v \\ 1 \end{bmatrix}\big) = \big(\frac{1-uv}{u-v}, i\frac{1+uv}{u-v}, \frac{u+v}{u-v}\big)$. A calculation shows that $\tilde{\sigma}(S^2) = \big\{(u,v) \mid v = -1/\bar{u}\big\} = \mathbf{C}P^1$, and in this way $\tilde{\sigma}\,|_{S^2} = \sigma$. Thus $A \mapsto \tilde{\sigma}^{-1} \circ A \circ \tilde{\sigma}$ extends $\Sigma$, and consequently we denote them by the same name

$$\begin{aligned} \Sigma \,:\, SL_2(\mathbf{C}) &\longrightarrow SO_3(\mathbf{C}) \\ A &\longmapsto \tilde{\sigma}^{-1} \circ A \circ \tilde{\sigma} \end{aligned}$$

Explicitly (see [12]) $\Sigma$ is given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \overset{\Sigma}{\longmapsto} \begin{pmatrix} \frac{1}{2}(a^2 - b^2 - c^2 + d^2) & \frac{i}{2}(a^2 + b^2 - c^2 - d^2) & cd - ab \\ \frac{i}{2}(d^2 + b^2 - c^2 - a^2) & \frac{1}{2}(a^2 + b^2 + c^2 + d^2) & i(ab + cd) \\ bd - ac & -i(bd + ac) & ad + bc \end{pmatrix}$$

This map is a Lie group homomorphism by construction, 2:1 by inspection, and thus an open covering map, since the induced isomorphism of Lie algebras $\Sigma_*$ is invertible at every point:

$$
\begin{aligned}
\Sigma_* \ : \ \mathfrak{sl}_2(\mathbf{C}) \quad &\longrightarrow \quad \mathfrak{so}_3(\mathbf{C}) \\
\begin{pmatrix} x_1 & x_2 \\ x_3 & -x_1 \end{pmatrix} &\mapsto \begin{pmatrix} 0 & 2ix_1 & x_3 - x_2 \\ -2ix_1 & 0 & i(x_2 + x_3) \\ x_2 - x_3 & -i(x_2 + x_3) & 0 \end{pmatrix}
\end{aligned}
$$

The images of certain $SL_2(\mathbf{C})$ subgroups will be useful in our study:

$$
\begin{pmatrix} e^{iz} & 0 \\ 0 & e^{-iz} \end{pmatrix} \overset{\Sigma}{\longmapsto} \begin{pmatrix} \cos 2z & -\sin 2z & 0 \\ \sin 2z & \cos 2z & 0 \\ 0 & 0 & 1 \end{pmatrix}
$$

$$
\begin{pmatrix} \cos z & -\sin z \\ \sin z & \cos z \end{pmatrix} \overset{\Sigma}{\longmapsto} \begin{pmatrix} \cos 2z & 0 & -\sin 2z \\ 0 & 1 & 0 \\ \sin 2z & 0 & \cos 2z \end{pmatrix}
$$

$$
\begin{pmatrix} \cos z & i\sin z \\ i\sin z & \cos z \end{pmatrix} \overset{\Sigma}{\longmapsto} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos 2z & -\sin 2z \\ 0 & \sin 2z & \cos 2z \end{pmatrix}
$$

In particular, Goursat's imaginary rotations are simply images of positive diagonal elements of $SL_2(\mathbf{R})$: if $A = diag\,\{r, \frac{1}{r}\}$ with $r > 0$, say $r = e^y$, then $\Sigma(A)$ is an imaginary rotation by $2iy$. Thus $\Sigma$ identifies Goursat's decomposition of $SO_3(\mathbf{C})$ with the polar decomposition of $SL_2(\mathbf{C})$, $A = PU$, where the positive Hermitian matrix $P$ is then diagonalized $P = \tilde{U}D\tilde{U}^*$ by a special unitary $\tilde{U}$, and the result $A = \tilde{U}D\tilde{U}^*U$ is mapped into $SO_3(\mathbf{C})$ by $\Sigma$.

## 2.2 A characterization of the Goursat transform

We are now ready to describe our own results. As deformations go, the Goursat transform has the important advantage that, once the periods of the conjugate surface are known, no additional period problems will occur after the transformation. This is because $\mathbf{x}_O = A\mathbf{x} - B\mathbf{y}$ when $O = A + iB$, and thus, at worst, $\mathbf{x}_O$ will acquire the periods of both $\mathbf{x}$ and $\mathbf{y}$, its coordinate functions being linear combinations of those of $\mathbf{x}$ and $\mathbf{y}$. This is certainly *not* the case for the conformal transform, which tends to introduce periods in completely unpredictable ways. In fact, one finds that even the most innocent looking modifications to the Weierstrass data will wildly alter the surface and its periods, the conformal transform being exactly such a modification, $(g, \eta) \to (g, f(z)\eta)$.

It is natural to wonder how certain changes in Weierstrass data affect a surface, and considering how well behaved the Goursat transform is with respect to the period problem, one might ask: If $\mathbf{x}$ is given by Weierstrass data $(g, \eta)$, what is the Weierstrass data $(\tilde{g}, \tilde{\eta})$ for the transform $\tilde{\mathbf{x}} = \mathbf{x}_O$? Equivalently, if $\mathbf{x}$ is given by spinor fields $(p, q)$, what are the spinor fields $(\tilde{p}, \tilde{q})$ for $\mathbf{x}_O$? The answer naturally leads to an alternative description of Goursat's double cover of $SO_3(\mathbf{C})$ by $SL_2(\mathbf{C})$.[1]

**Lemma 18.** *The spinor map $\varphi$ obeys the equivariance law* $\varphi\left(A \begin{pmatrix} p \\ q \end{pmatrix}\right) = \Sigma(A)\,\varphi\begin{pmatrix} p \\ q \end{pmatrix}$.

*Proof.* This is an elementary calculation, which we record here to help gain familiarity with the map $\Sigma$. In fact, this equivariance is easily seen to extend to the general linear group $GL_2(\mathbf{C}) = \mathbf{C} \times SL_2(\mathbf{C})$ and conformal group $CO_3(\mathbf{C}) = \mathbf{C} \times SO_3(\mathbf{C})$ actions, but since the $\mathbf{C}$ factor amounts to a constant conformal transform, we will not emphasize this extension.

Recall that $SL_2(\mathbf{C})$ is generated by diagonal elements $\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$, parabolic elements $\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$, and the inversion $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Thus it suffices to prove the claim for these generators:

[1]Since Goursat was aware of the Weierstrass representation and conceived of $\Sigma$ for the purposes of investigating his transform on surfaces, it seems miraculous that he used $\tilde{\sigma}$ to define $\Sigma$ when the spinor map $\varphi$ serves the same purpose: the former being perhaps the more natural construction to extend $\sigma$, the latter so immediately relevant to the study of minimal surfaces.

*Case*1 (Diagonal):

$$\begin{aligned}
\varphi(\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix}) &= \begin{pmatrix} \frac{1}{2}(q^2/a^2 - a^2p^2) \\ \frac{i}{2}(q^2/a^2 + a^2p^2) \\ qp \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{4}(a^2 + a^{-2})(q^2 - p^2) - \frac{1}{4}(a^2 - a^{-2})(q^2 + p^2) \\ \frac{i}{4}(a^{-2} - a^2)(q^2 - p^2) + \frac{i}{4}(a^2 + a^{-2})(q^2 + p^2) \\ qp \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{2}(a^2 + a^{-2}) & \frac{i}{2}(a^2 - a^{-2}) & 0 \\ \frac{i}{2}(a^{-2} - a^2) & \frac{1}{2}(a^2 + a^{-2}) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2}(q^2 - p^2) \\ \frac{i}{2}(q^2 + p^2) \\ qp \end{pmatrix} \\
&= \quad \Sigma \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \; \varphi\begin{pmatrix} p \\ q \end{pmatrix}
\end{aligned}$$

*Case*1 (Parabolic):

$$\begin{aligned}
\varphi(\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix}) &= \begin{pmatrix} \frac{1}{2}(q^2 - (p + bq)^2) \\ \frac{i}{2}(q^2 + (p + bq)^2) \\ q(p + bq) \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{4}(2 - b^2)(q^2 - p^2) - \frac{1}{4}b^2(q^2 + p^2) - bpq \\ \frac{i}{4}b^2(q^2 - p^2) + \frac{i}{4}(2 + b^2)(q^2 + p^2) + ibqp \\ \frac{1}{2}b(q^2 - p^2) + \frac{1}{2}b(q^2 + p^2) + qp \end{pmatrix} \\
&= \begin{pmatrix} 1 - \frac{1}{2}b^2 & \frac{i}{2}b^2 & -b \\ \frac{i}{2}b^2 & 1 + \frac{1}{2}b^2 & ib \\ b & -ib & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2}(q^2 - p^2) \\ \frac{i}{2}(q^2 + p^2) \\ qp \end{pmatrix} \\
&= \quad \Sigma \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \; \varphi\begin{pmatrix} p \\ q \end{pmatrix}
\end{aligned}$$

*Case*1 (Inversion):

$$\begin{aligned}
\varphi(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix}) &= \begin{pmatrix} \frac{1}{2}(p^2 - q^2) \\ \frac{i}{2}(p^2 + q^2) \\ -qp \end{pmatrix} \\
&= \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{2}(q^2 - p^2) \\ \frac{i}{2}(q^2 + p^2) \\ qp \end{pmatrix} \quad = \quad \Sigma(J) \; \varphi\begin{pmatrix} p \\ q \end{pmatrix}
\end{aligned}$$

$\square$

Our use of the spinor map has shown, according to Thm 3, that if $\mathbf{x}$ is described by spinor fields $\binom{p}{q}$, then $\mathbf{x}_{\Sigma(A)}$ is exactly the surface described by spinor fields $A\binom{p}{q}$. This is indeed the central fact for our entire study, but let us pursue a more geometric characterization:

**Theorem 19.** *Let* $\mathbf{x}$ *be a minimal surface with Gauss map* $g$ *and Hopf differential* $II^{2,0}$*. The Goursat transform* $\mathbf{x}_{\Sigma(A)}$ *is exactly the surface with Gauss map* $\tilde{g} = A \cdot g$ *and Hopf differential* $\widetilde{II}^{2,0} = II^{2,0}$.

Here " $A\cdot$ " denotes linear fractional transformation (Möbius transformation) on the extended complex plane $\widetilde{\mathbf{C}} := \mathbf{C} \cup \{\infty\}$:

$$A \cdot z := \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az+b}{cz+d}$$

Also note that a minimal surface is indeed specified by its Gauss map and Hopf differential, since knowing $g$ and $II^{2,0} = \eta dg$ allows one to recover the Weierstrass data $(g, \eta)$.

*Proof.* The identification $\varsigma : \widetilde{\mathbf{C}} \xrightarrow{\simeq} \mathbf{C}P^1$ given by $\varsigma(z) := \begin{bmatrix} z \\ 1 \end{bmatrix}$ respects LFT and projective transformation, respectively: $\varsigma\,(A \cdot z) = \left[A\binom{z}{1}\right] = A\begin{bmatrix} z \\ 1 \end{bmatrix} = A\,\varsigma(z)$. Since the spinor map is homogeneous, $\varphi\left(\lambda\binom{p}{q}\right) = \lambda^2\,\varphi\binom{p}{q}$ for any $\lambda \in \mathbf{C}^*$,

$$\begin{array}{ccc} \mathbf{C}^2 & \xrightarrow{\ \varphi\ } & Q_1 \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi} \\ \mathbf{C}P^1 & \xrightarrow[\simeq]{\ \tilde{\varphi}\ } & \mathcal{P}Q_1 \end{array}$$

it descends to a (biholomorphic) map $\tilde{\varphi} : \mathbf{C}P^1 \to \mathcal{P}Q_1$ of projective varieties, where $\mathcal{P}Q_1 := \{\mathbf{z} \in Q_1\} / \mathbf{z} \sim \lambda\mathbf{z}$, and $\pi$ is the natural projection. This map (called the *Segre map*) identifies the variety $\mathcal{P}Q_1$ as the Riemann sphere, and the equivariance of the spinor map $\varphi$ also holds for $\tilde{\varphi}$. Thus the composition

$$\phi := \tilde{\varphi}^{-1} \circ \varsigma^{-1} \circ \pi : Q_1 \longrightarrow \widetilde{\mathbf{C}} \quad \text{satisfies} \quad \phi\big(\Sigma(A)\mathbf{z}\big) = A \cdot \phi(\mathbf{z}).$$

The importance of $\phi$ is that the Gauss map of $\mathbf{x}$ can now be expressed simply as $g = \phi(\mathbf{x}_z)$, where $z$ is a local complex coordinate. It follows that the Gauss map of $\mathbf{x}_{\Sigma(A)}$ is

$$\tilde{g} = \phi((\mathbf{x}_{\Sigma(A)})_z)) = \phi(\Sigma(A)\mathbf{x}_z) = A \cdot \phi(\mathbf{x}_z) = A \cdot g$$

Next, recall that the Hopf differential can be expressed as $II^{2,0} = qdp - pdq$ which can be rewritten $d\begin{pmatrix} p \\ q \end{pmatrix}^T J \begin{pmatrix} p \\ q \end{pmatrix}$, where again $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Note the following commutation property for $J$ in $SL_2(\mathbf{C})$ : $JA = (A^T)^{-1}J$. We then compute:

$$\widetilde{II}^{2,0} = d(A\begin{pmatrix} p \\ q \end{pmatrix})^T J(A\begin{pmatrix} p \\ q \end{pmatrix}) = d\begin{pmatrix} p \\ q \end{pmatrix}^T A^T (A^T)^{-1} J \begin{pmatrix} p \\ q \end{pmatrix} = d\begin{pmatrix} p \\ q \end{pmatrix}^T J \begin{pmatrix} p \\ q \end{pmatrix} = II^{2,0}$$

$\square$

CMC surfaces are characterized by the holomophicity of the Hopf differential. In the case of minimal surfaces, the most natural modification to $II^{2,0}$ that preserves this condition is the multiplication of $II^{2,0}$ by a non-vanishing holomorphic function $f(z)$, a process which leaves the Gauss map unchanged. This is what we have termed the conformal transform. In the case of minimal surfaces, this transform has a kind of "complimentary" transform. Recall that minimal surfaces are characterized by the conformality of Gauss map. The most natural modification to $g$ that preserves this condition is the composition of $g$ with a conformal map $A : S^2 \to S^2$, a process which leaves the Hopf differential unchanged. But the conformal transformations of $S^2$ are exactly the Möbius transformations[2] $PSL_2(\mathbf{C}) = SL_2(\mathbf{C})/\{\pm I\}$ acting by linear fractional transformation. This is none other than the Goursat transform!

[2]In the sequel, we will almost always suppress the distinction between $SL_2(\mathbf{C})$ and $PSL_2(\mathbf{C})$, since for our purposes, this sign ambiguity should cause no confusion.

Not only does this characterization elucidate the transform's geometric meaning, it also seems to make it more manageable, in the sense that the $SL_2(\mathbf{C})$ action on spinor fields (or Möbius transformation of the Gauss map) is somehow easier to deal with and visualize than the action of $SO_3(\mathbf{C})$ on minimal curves. This will be helpful for a number of applications to important classes of minimal surfaces.

## 2.3 Complete minimal surfaces

Recall that the complete surfaces form a natural class of interest in the theory. Osserman's investigation of this class particularly addressed the possible size of the set of points omitted by the Gauss map, in the form of "Picard-like" theorems. We will often refer to this set as "the normal directions omitted by $\mathbf{x}$" or simply "the points omitted by $\mathbf{x}$." Many years of research, culminating in the work of Fujimoto [16], produced the optimal version of Osserman's result: a complete minimal surface can omit at most four points. Recently there have been efforts to ascertain the extent to which the set of omitted points can be "prescribed" in advance, that is, can we produce interesting examples of complete surfaces omitting a fixed set of points $\{z_1, \ldots, z_n\} \in S^2$? This is typically done by first choosing $M = S - \{p_1, \ldots, p_m\}$ (where $S$ is a compact Riemann surface) and a holomorphic map $g : S \to \widetilde{\mathbf{C}}$, then arguing that $\eta$ can be chosen in such a way that the induced metric is complete. Our results above suggest that the Goursat transform is uniquely suited to assist with this sort of construction, with the extra feature that we obtain new surfaces via *continuous deformations* of existing examples. The main result is that at least three omitted points of a complete surface can be prescribed.

**Lemma 20.** *Let $\mathbf{x} : M^2 \to \mathbf{R}^3$ be a complete (resp. non-complete) minimal surface. Then any Goursat transform $\mathbf{x}_O$ is also complete (resp. non-complete).*

*Proof.* If $\mathbf{x}$ is given by spinor fields $(p, q)$, then the induced metric can be expressed

$$ds^2 = \left\| \begin{pmatrix} p \\ q \end{pmatrix} \right\|^4$$

Let $d\tilde{s}^2$ denote the metric of $\mathbf{x}_O$ and choose $A \in SL_2(\mathbf{C})$ such that $O = \Sigma(A)$. Since $A^*A$ is Hermitian, it can be diagonalized by a unitary matrix, say $A^*A = U^*\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}U$. Then

$$\left\| A\begin{pmatrix} p \\ q \end{pmatrix} \right\|^2 = \begin{pmatrix} p \\ q \end{pmatrix}^* A^*A\begin{pmatrix} p \\ q \end{pmatrix} = \left(U\begin{pmatrix} p \\ q \end{pmatrix}\right)^*\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\left(U\begin{pmatrix} p \\ q \end{pmatrix}\right).$$

Assuming that $\lambda_1 \leq \lambda_2$, we have

$$\lambda_1 \left\| \begin{pmatrix} p \\ q \end{pmatrix} \right\|^2 \leq \left(U\begin{pmatrix} p \\ q \end{pmatrix}\right)^*\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\left(U\begin{pmatrix} p \\ q \end{pmatrix}\right) \leq \lambda_2 \left\| \begin{pmatrix} p \\ q \end{pmatrix} \right\|^2$$

$$\Longrightarrow \quad \lambda_1^2\, ds^2 \leq d\tilde{s}^2 \leq \lambda_2^2\, ds^2$$

Thus curves of finite (resp. infinite) length in the original surface will have finite (resp. infinite) length in the transform, so completeness (resp. non-completeness) is preserved. □

*Remark* 21. Of course, the same argument using the original $SO_3(\mathbf{C})$ action could have established the result. We prefer this demonstration because the eigenvalues of $A^*A$ are easily determined for 2×2 matrices, which provide quick estimates on the distortion of distances. These eigenvalues are particularly helpful for the usual generators of $SL_2(\mathbf{C})$, which are automatic for diagonal and off-diagonal elements, and $\lambda_{1,2} = \frac{1}{2}(r^2 + 2 \pm r\sqrt{r^2+4})$ for a parabolic element $\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$ with $b = re^{i\theta}$.

**Theorem 22.** *Let* $\tilde{\mathbf{x}} : M^2 \to \mathbf{R}^3$ *be a complete minimal surface whose Gauss map omits* $n$ *points* $\{\tilde{z}_1, \ldots, \tilde{z}_n\}$*. Then* $\tilde{\mathbf{x}}$ *can be continuously deformed to a complete* *surface* $\mathbf{x}$ *whose Gauss map omits* $n$ *points* $\{z_1, \ldots, z_n\}$*, three of which* $\{z_1, z_2, z_3\}$ *may be prescribed.*

*Proof.* A Möbius transformation of the Riemann sphere is bijective, so certainly the number of omitted points $n$ will be the same for any Goursat transform. Furthermore, the action of the Möbius group is strictly three-transitive, so there exists a unique $A \in SL_2(\mathbf{C})$ such that $A \cdot \tilde{z}_i = z_i$ for $1 \leq i \leq 3$. Letting $a(t)$ be a smooth path $a : \mathbf{R} \to SL_2(\mathbf{C})$ connecting $A$ to the identity $I$, say $a(0) = I, a(1) = A$, then by the above theorem $\tilde{\mathbf{x}}_{\Sigma(a(t))}$ is the desired deformation, with $\mathbf{x} = \tilde{\mathbf{x}}_{\Sigma(a(1))}$. $\square$

As suggested in the proof of Theorem 7, the case $n = 3$ is particularly interesting. A triple of points $\{z_1, z_2, z_3\}$ specifies a Möbius transformation, namely, the unique $A$ such that $A : \{z_1, z_2, z_3\} \mapsto \{0, 1, \infty\}$, and conversely, any Möbius transformation $A$ defines a unique triple $\{z_1, z_2, z_3\} := A^{-1}\{0, 1, \infty\}$. Thus we have

**Corollary 23.** *Let* $\mathbf{x}$ *be a complete minimal surface omitting three normal directions. Then the set of Goursat transforms* $\{\mathbf{x}_O\}$ *is in* $1-1$ *correspondence with triples of points in* $S^2$*, namely, each surface omits a unique triple, and each triple is omitted by a unique surface.*

Now let us narrow our study to Osserman's class of complete FTC surfaces. We will show that the Goursat transform preserves a certain subclass of these surfaces. Let $\mathbf{x}$ have associated minimal curve $\boldsymbol{\gamma}$. Recall that the "period vectors" of $\mathbf{x}$ are $\mathbf{v}_i = \oint_{\gamma_i} \dot{\boldsymbol{\gamma}}\, ds$, where the $\gamma_i$ are generators of the fundamental group $\pi_1(M^2)$. Then $\mathbf{x}$ descends to $M^2$ if $\mathbf{v}_i \in i\mathbf{R}^3\ \forall i$ and its conjugate $\mathbf{y}$ descends if $\mathbf{v}_i \in \mathbf{R}^3\ \forall i$. We are interested in the case $\mathbf{v}_i = \mathbf{0}$, so that *both* surfaces descend.

**Corollary 24.** *Let* $\mathbf{x} : M^2 \to \mathbf{R}^3$ *be a complete FTC minimal surface with vanishing period vectors. Then any Goursat transform* $\mathbf{x}_O$ *is also complete FTC. Moreover, the total curvature is invariant under the transformation.*

*Proof.* By Osserman's theorem, $M^2$ is conformally equivalent to a finitely punctured compact Riemann surface $S-\{p_1,\ldots,p_n\}$ and the Gauss map $g$ extends holomorphically to all of $S$. Conversely, in the event that a minimal immersion of $S-\{p_1,\ldots,p_n\}$ has no real periods and its Gauss map $\tilde{g}$ is well-defined on all of $S$, then the total curvature is $-4\pi$ times the degree of the Gauss map. Our hypotheses imply that this is the case for $\mathbf{x}_O$ as well, since $\mathbf{x}_O = A\mathbf{x} - B\mathbf{y}$ cannot have real periods, and by our characterization of the Goursat transform, $\tilde{g} = A \cdot g$. Since a Möbius transformation is bijective, $\mathbf{x}_O$ has total curvature

$$\iint\limits_{M^2} \tilde{K}\, dA \; = \; 4\pi\, deg(A\cdot g) \; = \; 4\pi\, deg(g) \; = \; \iint\limits_{M^2} K\, dA \qquad \square$$

If indeed the Goursat transform is to be regarded as a continuous deformation, this result should be expected, since the total curvature of complete minimal surfaces is quantized. Again, the case $n = 3$ is very interesting to consider, but surprisingly (to the author's knowledge, at the time of this writing) there is no known example of a complete FTC whose Gauss map omits exactly three points. Thus it is not known if $n = 3$ is a sharp upper-bound for Osserman's theorem in the FTC case! An example of a complete FTC omitting three normal directions with vanishing periods becomes especially desirable for us since it would again give rise to a family of complete FTC surfaces in bijection with triples of points in $S^2$.

However, the reader with some experience in producing FTC minimal surfaces might recognize a certain naivety in considering only period-free examples, the real period problem alone often proving insurmountable. It is still possible to salvage our strategy for surfaces with "almost" vanishing periods:

**Definition 25.** A minimal surface has *parallel periods* if its period vectors $\mathbf{v}_i$ span a real line in $\mathbf{C}^3$.

The Catenoid is an example of a complete surface with parallel periods and finite total curvature. Recall that after a rotation, each matrix $O \in SO_3(\mathbf{C})$ has (up to sign) a unique real unit eigenvector $\mathbf{v}$. If $\mathbf{x}$ has periods parallel to $\mathbf{v}$, then the transform $\mathbf{x}_O$ will (after rotation) also have periods parallel to $\mathbf{v}$. We can use this fact to slightly generalize the result in the FTC case, albeit with quite a bit of added restriction:

**Proposition 26.** *Let $\tilde{\mathbf{x}} : M^2 \to \mathbf{R}^3$ be a complete FTC surface with parallel periods, omitting at least one pair of non-antipodal points $\{\tilde{z}_1, \tilde{z}_2\}$. Then $\tilde{\mathbf{x}}$ can be continuously deformed to a complete FTC surface $\mathbf{x}$ whose Gauss map omits any prescribed pair of points $\{z_1, z_2\}$.*

*Proof.* Since $\tilde{\mathbf{x}}$ is FTC, its period vectors $\mathbf{v}_i$ must be parallel to $i\mathbf{v}$ where $\mathbf{v}$ is a unit vector in $\mathbf{R}^3$. If $O\mathbf{v} = \mathbf{v}$, then the period vectors $O\mathbf{v}_i$ of $\mathbf{x}_O$ are also imaginary, and thus $\mathbf{x}_O$ is complete FTC. The subgroup $\{A \in SL_2(\mathbf{C}) \mid \Sigma(A)\mathbf{v} = \mathbf{v}\}$ is isomorphic to $\{A \in SL_2(\mathbf{C}) \mid \Sigma(A)\mathbf{k} = \mathbf{k}\}$ where $\mathbf{k} = (0, 0, 1)$, so by rotation we may assume that $\mathbf{v} = \mathbf{k}$. Recall from the first section that $\{A \in SL_2(\mathbf{C}) \mid \Sigma(A)\mathbf{k} = \mathbf{k}\}$ is exactly the diagonal matrices, which include dilations of $\widetilde{\mathbf{C}}$ (real diagonals). A pair of non-antipodal points $\{\tilde{z}_1, \tilde{z}_2\}$ can be moved arbitrarily close by the continuous action of this subgroup, so there exists a curve $A : [0, 1] \to SL_2(\mathbf{C})$ such that $A_0 = I$ and $dist\,(A_1 \cdot \tilde{z}_1, A_1 \cdot \tilde{z}_2) = dist\,(z_1, z_2)$. If $U$ is the unitary accomplishing $UA_1 \cdot \tilde{z}_i = z_i$, then $\mathbf{x}_{\Sigma(UA_t)}$ is the desired deformation. □

## 2.4 Examples

When a minimal surface has finite topology, $M^2 \simeq S - \{p_1, ..., p_n\}$, the points omitted by the Gauss map $g$ are typically the images $\tilde{g}(p_i)$, where $\tilde{g}$ is the (meromorphic) extension of $g$ to $S$. Then we can expect to visualize these omitted points by looking at the position of the ends, since the ends occur at the $p_i$. For example, the Catenoid omits 0 and $\infty$, corresponding to the fact that one end points vertically downward (the end at 0) and the other end points vertically upward (the end at $\infty$). End positions are key features of the geometry of a surface, so this phenomena provides an easy strategy to visualize the above theorems.

*Example* 2. (Catenoid Bending):

The Catenoid, defined by Weierstrass data $(\frac{1}{z}, dz)$, is a complete surface defined on the punctured plane, $M^2 = \mathbf{C} - \mathbf{0}$, which has the maximal non-zero total curvature, $-4\pi$. Unfortunately, it is not subject to Prop 26, since its omitted points 0 and $\infty$ are antipodal, but we can still obtain interesting complete periodic examples with *helicoidal* ends. Parabolic Goursat transformations by $\Sigma(A)$ with $A = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$ effect $A : 0 \mapsto b$ and $A : \infty \mapsto \infty$ fixed. Since the omitted points are just the directions of the ends, such a transform forces the bottom end to "bend" towards the top.

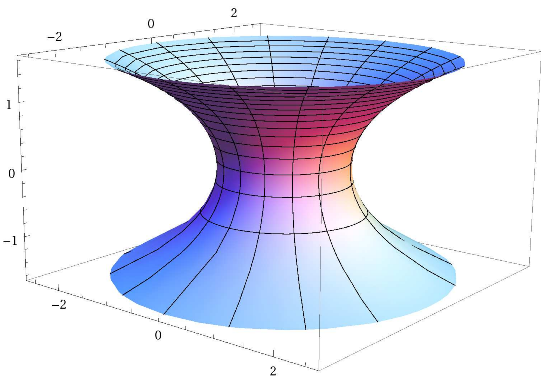

**Fig.2.1** $b = 0$

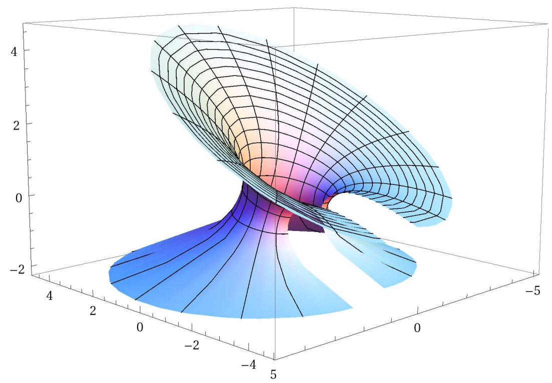

**Fig.2.2** $b = 1/2$

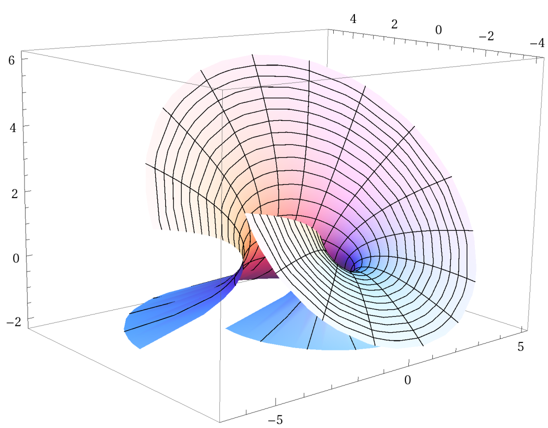

**Fig.2.3** $b = \frac{i}{2}$

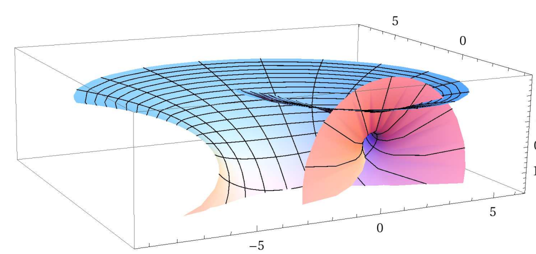

**Fig.2.4** $b = i$

These graphics show the surfaces upside-down, so now the *bottom* end (at $\infty$) stays fixed, while the top end (at 0) bends to point in the direction of $b$. Notice that the ends are orthogonal in the last image, one end points along the positive z-axis (toward $\infty$), the other along the negative x-axis (toward $i$) These transforms all pick up periods from the conjugate (the Helicoid), which is responsible for the visible self-intersection in the lower two images. Since parabolic transformations can adjust the distance between the omitted points arbitrarily, any prescribed pair of omitted points can be obtained in this way, after an appropriate rotation. A diagonal transformation fixes both of the original omitted points, and by a change of variables such a transform is seen to be a reparameterization of the original surface. The fact that the Catenoid avoids Prop 26 suggests a kind of "Catenoidal rigidity" theorem[3]:

*Remark* 27. A complete embedded minimal surface of maximal non-zero total curvature omits two points which must be antipodal.

*Example* 3. (Voss surface):

Voss [22] noticed a simple construction of complete surfaces which omit any (allowable) number of points. Let $M^2 = S^2 - \{p_0, ...p_n\}$, and choose $p_0 = \infty$ so that $M^2 \simeq \mathbf{C} - \{p_1, ...p_n\}$. The minimal immersion with data $( z, \prod_{i=1}^{n}(z-p_i)^{-1}dz)$ clearly omits $n+1$ points, and a calculation shows that the induced metric is complete if and only if $n \leq 3$. To apply our theorem, let us take $n = 2$ and set $\{p_1, p_2\} = \{1, -1\}$, which we call the *Voss surface*. It has a saddle-shaped central piece:

[3]This is indeed a trivial consequence of the actual, and far more interesting, classification result of Osserman [16], that the Catenoid is the unique complete embedded minimal surface of total curvature $-4\pi$.

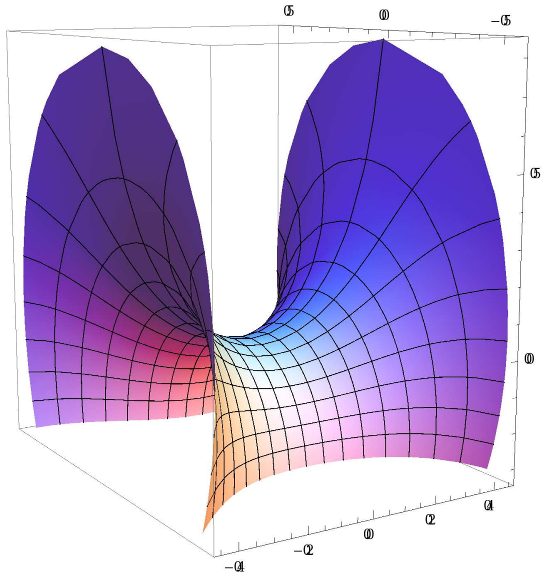

**Fig.2.5** Voss Surface

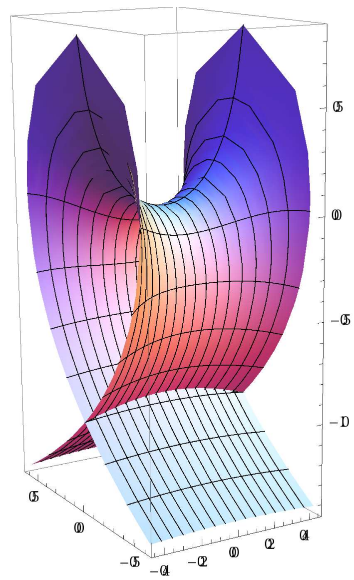

**Fig.2.6** Larger domain

The sides of the saddle that curve upward are the ends at 1 and $-1$, and the sides that curve downward (where they eventually intersect, as seen in the second image) form the end at infinity. It has one plane of symmetry passing through (inverse stereographic projections of) the three omitted points $1, -1, \infty$. Diagonal Goursat transformations by $\Sigma(A)$ with $A = \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$ effect $A : 1 \mapsto a^2$, $A : -1 \mapsto -a^2$

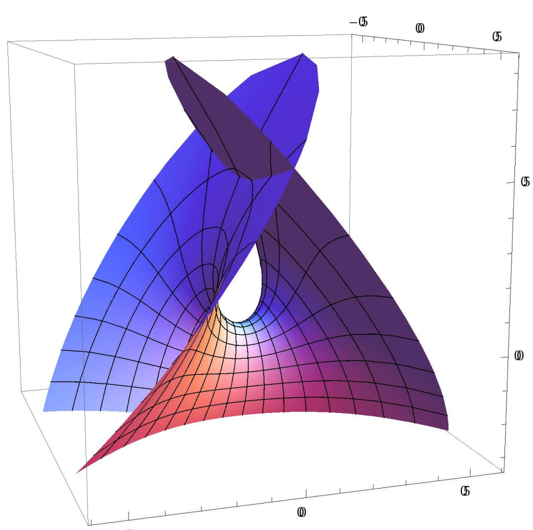

**Fig.2.7** $a = 1 + i$

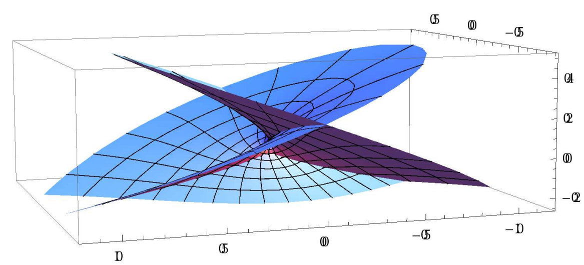

**Fig.2.8** $a = 2 + 2i$

Again we use use our theorem to interpret these images: the sides are forced to angle themselves towards the new directions $\pm a^2$, while the position of the infinity end is preserved, and as $a$ becomes larger, the sides become closer to parallel and the surface appears flatter. These transforms continue to have vertical symmetry (the symmetry plane is rotated by $2arg(a)$), which can be tilted by the presence of off-diagonal terms, for example, $A = \begin{pmatrix} 2 & 0 \\ 1 & \frac{1}{2} \end{pmatrix}$:

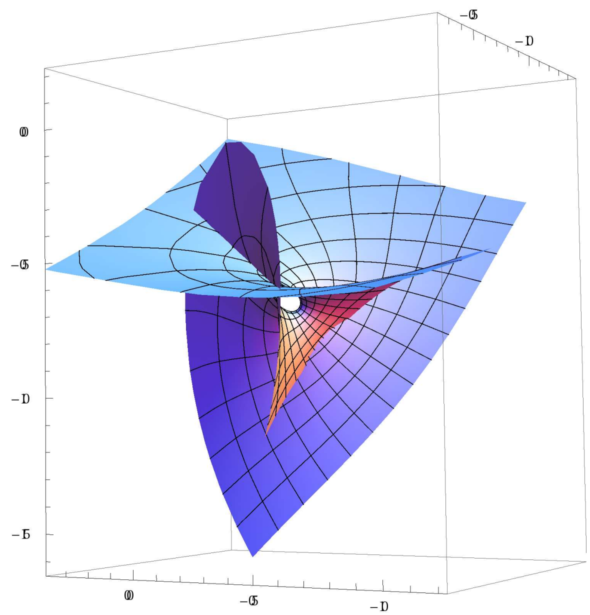

**Fig.2.9**

As noted above, the family of transforms $\mathbf{x}_{\Sigma(A)}$ of the Voss surface is in one-to-one correspondence with the set of triples in $S^2$. In this case, those triples are exactly the (normal) directions of the three ends, so the correspondence can be seen directly from the examples by looking at end positions. Of course, simply repeating the Voss construction for various choices of triples $\{p_0, p_1, p_2\}$ would constitute a family with the same properties, without resorting to this deformation. The advantage of using the Goursat transform will become clear in the next chapter.

*Example* 4. (B-surfaces):

Let $M^2 = \mathbf{C} - 0$ and consider the minimal immersions defined by data $(\frac{z^2}{k}, \frac{k\,dz}{z})$ where $k$ is an arbitrary non-zero complex constant. These are complete surfaces omitting 0 and $\infty$ at its two ends, which we call the *Bessel surface*, for reasons we explain later. We will focus on one part of the surface which consists of a central piece with two catenoidal-necks branching off. Unlike the Catenoid, these necks are not "closed": each neck self-intersects and extends out in two directions to form a third lower neck. The necks and their intersecting lips can be seen by varying the parameter domain, and we note at least one obvious vertical plane of symmetry:

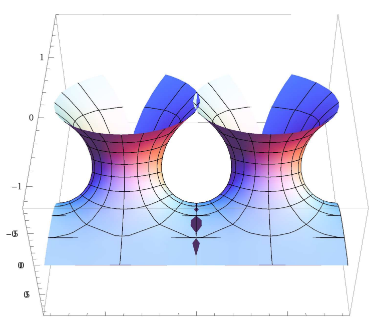

**Fig.2.10** Above

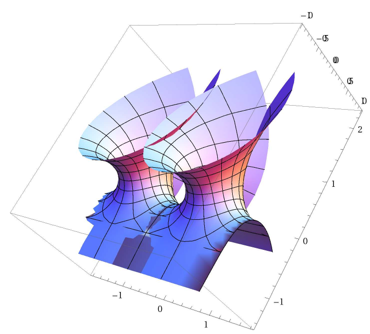

**Fig.2.11** Side

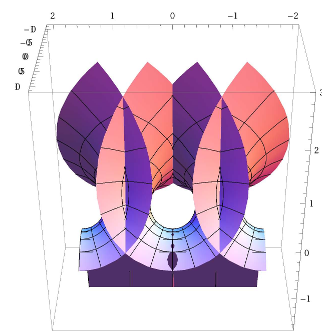

**Fig.2.12** Below

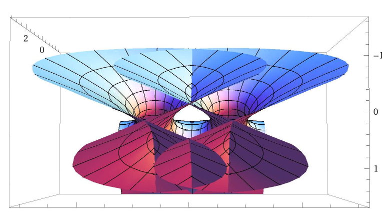

**Fig.2.13** Front

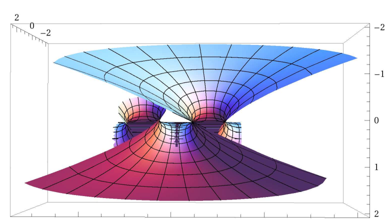

**Fig.2.14** Front

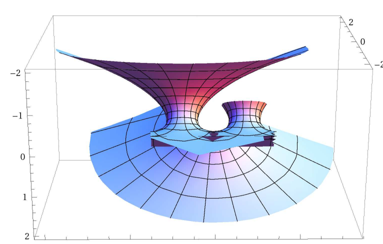

**Fig.2.15** Back

This last image only displays the extension of one of the necks, allowing us to see the interior more clearly. It is important to note that the necks all represent the *same* end, namely the end at infinity, and the central connecting piece is the end at zero.

Again consider parabolic Goursat transformations by $\Sigma(A)$ with $A = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$. Thus the direction of the necks must stay fixed (since $A$ fixes infinity) and the connecting base must turn so that it faces the $b$-direction. The following graphics use the same parameter domain as in the last figure:

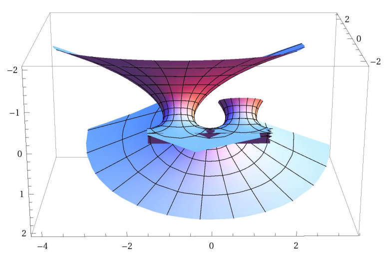

**Fig.2.16** $b = 0$

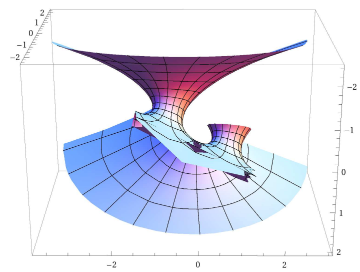

**Fig.2.17** $b = \frac{i}{4}$

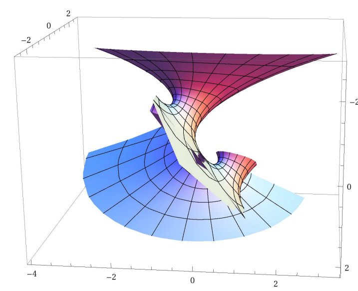

**Fig.2.18** $b = \frac{i}{2}$

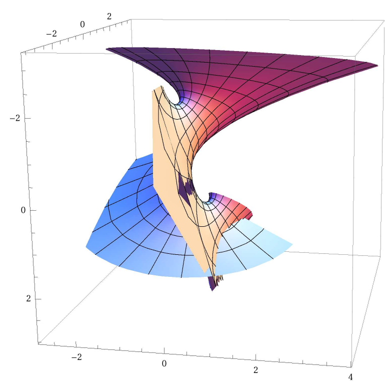

**Fig.2.19** $b = \frac{3i}{4}$

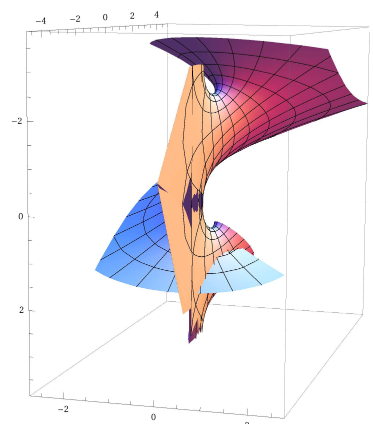

**Fig.2.20** $b = i$

We indeed observe the predicted effect: the base piece rotates through an angle of $\pi$ as we adjust $b$ from 0 to $i$.

## 2.5 Cartan's invariant

As we have seen, the Goursat geometry of minimal surfaces (i.e. that geometry[4] with group $E_3(\mathbf{C})$ acting on the space of minimal surfaces) generalizes the usual Euclidean geometry (with group $E_3(\mathbf{R}) \subset E_3(\mathbf{C})$), and by Corollary 4, the Schwarzian derivative of the Gauss map $\mathcal{S}_z(g)$ and the Hopf differential $II^{2,0}$ form a complete set of invariants for this geometry.

But, after all, minimal surfaces are Goursat-equivalent if and only if their associated minimal curves are $E_3(\mathbf{C})$-equivalent. Thus to find invariants for minimal surfaces, one could instead analyze minimal curves to produce a notion of "curvature," and in fact E. Cartan took exactly this approach in his classic text [6]. The obvious way to do so would be to construct a Frenet frame along the curve, following the derivation of curvature for the case of real curves in $\mathbf{R}^3$. However, the velocity vector of a minimal curve is null, and such a vector cannot be made into the first column of an $SO_3(\mathbf{C})$-frame, so Frenet's construction does not go through. Cartan nevertheless found a very clever way to produce an analogue of the Frenet frame despite this difficulty. He transforms an $SO_3(\mathbf{C})$-frame $F$ along a curve $\boldsymbol{\gamma}$ into a different type of $GL_3(\mathbf{C})$-frame, which he calls a *direct cyclic frame*, via $F \mapsto FE$, where $E$ is the matrix with column vectors $E_1 = (e_1 + i\, e_2)/2$, $E_2 = e_3$, $E_3 = e_1 - i\, e_2$, and $\{e_i\}$ is the standard basis of $\mathbf{C}^3$. He then carries out the usual reduction procedure on these new frames, thereby obtaining notions of arclength and curvature.

The curvature thus obtained is indeed a useful invariant for minimal surfaces, but one might complain that the solution of the problem is not particularly intuitive, nor is this invariant easily interpreted in any clear way in terms of the geometry of minimal

[4]Here we are using the term "geometry" in a generalized sense of Klein: a *geometry* is a pair $(X, G)$, where $G$ is a group (typically a Lie group) acting on a set $X$ (typically a manifold, moduli space of manifolds, or bundle over a manifold).

surfaces, or even that of the curves themselves. We now describe an alternative spin-geometric approach which helps address these issues. Observe that each minimal curve $\boldsymbol{\gamma}$ has (up to an additive constant) two associated curves $\pm\boldsymbol{\psi} : M^2 \to \mathbf{C}^2$ such that $d\boldsymbol{\gamma} = \varphi\,(\boldsymbol{\psi}')dz$. Namely, if $\boldsymbol{\gamma}$ is represented by spinors $(p, q)$, then the antiderivative

$$\boldsymbol{\psi}(z) = \int \begin{pmatrix} p \\ q \end{pmatrix} dz \;:\; M^2 \longrightarrow \mathbf{C}^2$$

is (up to sign, plus a constant) the unique holomorphic curve such that $d\boldsymbol{\gamma} = \varphi\,(\boldsymbol{\psi}')dz$. Conversely, a regular holomorphic curve $\boldsymbol{\psi} : M^2 \to \mathbf{C}^2$ gives rise to a minimal curve $\boldsymbol{\gamma} = \int \varphi(\boldsymbol{\psi}')dz$, and $\pm\boldsymbol{\psi} + \mathbf{v}$, with $\mathbf{v} \in \mathbf{C}^2$ constant, are the only curves that produce this $\boldsymbol{\gamma}$. By the equivariance of $\varphi$, a complex rotation of a minimal curve $\boldsymbol{\gamma}$ by $\Sigma(A)$ is equivalent to the affine transformation of its "spin curve" $\boldsymbol{\psi}$ by $A$. Thus, associating curves in this way, we might reasonably regard the special affine geometry of regular holomorphic curves $\boldsymbol{\psi} : M^2 \to \mathbf{C}^2$ as the *spin geometry*[5] of minimal curves $\boldsymbol{\gamma} : M^2 \to \mathbf{C}^3$ under complex rotation. Then to carry out the frame reduction on a minimal curve $\boldsymbol{\gamma}$ with $SO_3(\mathbf{C})$-frame $F$, instead of changing to a direct cyclic frame, we can lift *both* the curve and the frame to a spin curve $\boldsymbol{\psi}$ and an $SL_2(\mathbf{C})$-frame $\widetilde{F}$, which can be analyzed in the usual way since, in the spin setting, there is no obstruction to the Frenet construction.

Not surprisingly, this strategy recovers Cartan's original expression for curvature, and in fact, the most natural Frenet frames turn out to be the exact analogue of Cartan's cyclic Frenet frames. The main advantage to the spin approach is that, once the equivalence of Cartan's curvature with affine curvature has been established, we

[5] Again, we use the term in a very general (Kleinian) sense: the *spin* of a geometry $(X, G)$ with a non-simply connected group, is a geometry $(\widetilde{X}, \widetilde{G})$, where $\pi : \widetilde{G} \to G$ is the universal covering group of $G$, and $\widetilde{X}$ is some appropriate analogue of $X$ (for example, if $X$ is the tangent bundle, $\widetilde{X}$ is the spinor bundle). Looking for interesting geometric phenomena in the spin setting, and the process of lifting geometric problems to their spin analogues, appears to be one of the fundamental developments of late twentieth-century Riemannian geometry.

can use the classical theory of affine curves to dramatically simplify the form of the invariants. This simplification provides a means to interpret the invariants in a more geometric way and illuminates the relationship between the geometry of a minimal surface and that of its spin curve.

The following moving frame argument is strongly based on the exposition in Jensen [12]. $G = \mathbf{C}^2 \rtimes SL_2(\mathbf{C})$ acts transitively on $\mathbf{C}^2$ according to $(\mathbf{v}, A) \cdot \boldsymbol{\psi} := \mathbf{v} + A\boldsymbol{\psi}$, and the isotropy subgroup at the origin $\mathbf{0}$ is $SL_2(\mathbf{C})$. Identifying $\mathbf{C}^2$ with its image under the inclusion $\mathbf{C}^2 \hookrightarrow \mathbf{C}^3$, $\mathbf{v} \mapsto \begin{pmatrix} 1 \\ \mathbf{v} \end{pmatrix}$, the matrix representation $G \to SL_3(\mathbf{C})$

$$(\mathbf{v}, A) \longmapsto \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{v} & A \end{pmatrix}$$

realizes this action as a linear action. We denote the natural projection by $\pi : G \to \mathbf{C}^2$, $\pi(\mathbf{v}, A) = \mathbf{v}$, and define a *frame* along an immersion $\boldsymbol{\psi} : M^2 \to \mathbf{C}^2$ is a lift $F : M^2 \to G$. If $F$ is a fixed frame along $\boldsymbol{\psi}$, any other frame is given by $\widetilde{F} = FA$, where $A : M^2 \to G_0$,

$$G_0 = \left\{ \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & A \end{pmatrix} \;\middle|\; A \in SL_2(\mathbf{C}) \right\}$$

Let $\omega$ denote the Maurer-Cartan 1-form on $G$. By abuse of notation, we also use $\omega$ to denote the pull-back of the Maurer-Cartan form $F^{-1}dF$ by the frame $F$, which in our matrix representation can be expressed

$$\omega = F^{-1}dF = \begin{pmatrix} 0 & 0 & 0 \\ \omega^1 & \omega_1^1 & \omega_2^1 \\ \omega^2 & \omega_1^2 & -\omega_1^1 \end{pmatrix}$$

i.e. if $F = (\boldsymbol{\psi}, A)$ and $\mathbf{e}_i$ denote the column vectors of $A$, then $d\boldsymbol{\psi} = \omega^i \mathbf{e}_i$ and $d\mathbf{e}_j = \omega_j^i \mathbf{e}_i$. We define a frame $F$ to be *first order* if $\omega^2 = 0$. Geometrically, this simply says that $\mathbf{e}_1$ is a holomorphic multiple of the tangent vector $\boldsymbol{\psi}'$. Consequently, if $F$ is first order, any other first order frame is given by $\widetilde{F} = FA$ where $A : M^2 \to G_1$,

$$G_1 = \left\{ \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & A \end{pmatrix} \;\middle|\; A = \begin{pmatrix} a & b \\ 0 & a^{-1} \end{pmatrix} \right\}$$

which changes the Maurer-Cartan form according to $\tilde{\omega} = \widetilde{F}^{-1} d\widetilde{F} = A^{-1}\omega A + A^{-1}dA =$

$$\begin{pmatrix} 0 & 0 & 0 \\ a^{-1}\omega^1 & \omega_1^1 - ab\,\omega_1^2 & a^{-2}\omega_2^1 + 2a^{-1}b\,\omega_1^1 - b^2\omega_1^2 \\ 0 & a^2\omega_1^2 & -\omega_1^1 + ab\,\omega_2^1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & a^{-1}a' & a^{-1}b' + a^{-2}b\,a' \\ 0 & 0 & -a^{-1}a' \end{pmatrix} dz$$

where $(U, z)$ is a local complex coordinate. We will assume that $\omega_1^2 \neq 0$. Then $a$ can be chosen so that $\tilde{\omega}^1 = \tilde{\omega}_1^2$. This is the next step in the reduction: a first order frame $F$ is called *second order* if $\omega^1 = \omega_1^2$. By the transformation law for the Maurer-Cartan form above, if $F$ is second order, any other second order frame is given by $\widetilde{F} = FA$ where $A : M^2 \to G_2$,

$$G_2 = \left\{ \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & A \end{pmatrix} \;\middle|\; A = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \right\}$$

Second order frames always exist on an open neighborhood of a point where $\omega_1^2 \neq 0$. The remaining freedom in choosing a second order frame $F \mapsto \widetilde{F} = FA$ accomplishes

$$\tilde{\omega} \;=\; \begin{pmatrix} 0 & 0 & 0 \\ \omega^1 & \omega_1^1 - b\,\omega^1 & \omega_2^1 + 2b\,\omega_1^1 - b^2\omega^1 + b'dz \\ 0 & \omega^1 & -\omega_1^1 + b\,\omega^1 \end{pmatrix}$$

A second order frame $F$ is called *third order* if $\omega_1^1 = 0$. This is the Frenet frame: clearly we can reduce a second order frame to a unique third order frame, which exhausts our freedom. Since $\omega^1 = \tilde{\omega}^1$, this is a globally well-defined 1-form, and the structure equations $d\omega = -\omega \wedge \omega$ imply that $d\omega^1 = 0$, so locally $\omega^1 = d\sigma$, where $\sigma : M^2 \to \mathbf{C}$ is a holomorphic function called the *psudeo-arc parameter*. Since $\omega^1$ spans $\bigwedge^{1,0}(M^2)$, we can write $\omega_2^1 = k\,\omega^1$, where $k : M^2 \to \mathbf{C}$ is a holomorphic function called the *curvature*.

To see that this is none other than Cartan's invariant, let $(U, z)$ be a local coordinate, and write $d\boldsymbol{\psi} = \begin{pmatrix} p(z) \\ q(z) \end{pmatrix} dz$. Consider the frame $F_1 = (\boldsymbol{\psi}, A)$, where $A = (\mathbf{e}_1, \mathbf{e}_2) = \begin{pmatrix} g & -1 \\ 1 & 0 \end{pmatrix}$ and $g = p/q$. This frame[6] is defined on all of $M^2$ except a discrete set of points (the zero locus of $q$), and is first order, since $d\boldsymbol{\psi} = q\,\mathbf{e}_1$ implies that $\omega^1 = q(z)\,dz$, $\omega^2 = 0$. Since $d\mathbf{e}_1 = -dg\,\mathbf{e}_2$ and $d\mathbf{e}_2 = 0$, we have that $\omega^1_1 = \omega^1_2 = 0$ and $\omega^2_1 = -dg$:

$$\omega \;=\; F_1^{-1} dF_1 \;=\; \begin{pmatrix} 0 & 0 & 0 \\ q\,dz & 0 & 0 \\ 0 & -dg & 0 \end{pmatrix} \;=\; \begin{pmatrix} 0 & 0 & 0 \\ q & 0 & 0 \\ 0 & -g' & 0 \end{pmatrix} dz$$

By the calculation above, the frame $F_2 = F_1 A$ with diagonal $A : M^2 \to G_1$ effects

$$A = \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \quad \Longrightarrow \quad \tilde{\omega} \;=\; F_2^{-1} dF_2 \;=\; \begin{pmatrix} 0 & 0 & 0 \\ a^{-1} q & a^{-1} a' & 0 \\ 0 & -a^2 g' & -a^{-1} a' \end{pmatrix} dz$$

and setting $a = q(p\,q' - qp')^{-1/3}$ puts $\tilde{\omega}^1 = \tilde{\omega}^2_1 = (p\,q' - qp')^{1/3} dz = a^{-1} q\,dz$, making $F_2$ second order.[7] Finally, if $F_3 = F_2 A$ with $A : M^2 \to G_2$, then

$$A = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \quad \Longrightarrow \quad \omega \;=\; F_3^{-1} dF_3 \;=\; \begin{pmatrix} 0 & 0 & 0 \\ \tilde{\omega}^1 & \tilde{\omega}^1_1 - b\,\tilde{\omega}^1 & 2b\,\tilde{\omega}^1_1 - b^2 \tilde{\omega}^1 + b' dz \\ 0 & \tilde{\omega}^1 & -\tilde{\omega}^1_1 + b\,\tilde{\omega}^1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 \\ a^{-1} q & a^{-1} a' - b\,a^{-1} q & 2b\,a^{-1} a' - b^2 a^{-1} q + b' \\ 0 & a^{-1} q & -a^{-1} a' + b\,a^{-1} q \end{pmatrix} dz$$

Thus $b = a' q^{-1}$ accomplishes $\omega^1_1 = 0$, so that $F_3$ is third order. Finally, $\omega^1 = a^{-1} q\,dz$ and $\omega^1_2 = (aq)^{-1} (a')^2 + (a' q^{-1})' dz$ imply that

$$k = q^{-2} (a')^2 + a\,q^{-1} (a' q^{-1})'$$

---

[6] Compare $(\boldsymbol{\psi}, A)$ and its reduction to the direct cyclic frame $(\boldsymbol{\gamma}, O)$ built from the Weierstrass representation in section 8.3.10 of Jensen [] (there it is denoted $(\boldsymbol{\gamma}, A)$ instead of $(\boldsymbol{\gamma}, O)$) and verify that $OE^{-1} = \Sigma(A)$, that is, $(\boldsymbol{\psi}, A)$ is an $SL_2(\mathbf{C})$-lift of the $SO_3(\mathbf{C})$-frame $(\boldsymbol{\gamma}, OE^{-1})$.

[7] Note that the "speed" of the spin curve $\boldsymbol{\psi}$ differs from the speed of it's corresponding minimal curve $\boldsymbol{\gamma}$ by a power of 2/3.

Comparing this with the expression for Cartan's invariant $k$ [12] (and a severely long computation) shows that the two agree. Again, this is a helpful invariant for minimal surfaces since it distinguishes Goursat-equivalence classes. However, besides its complicated formula, its characteristics are difficult to relate to minimal surface geometry. For example, we it is not related to the curvature of surfaces: the *flat* surfaces ($k \equiv 0$ identically) are *not* planes, but in fact the (Goursat) similarities of the Catenoid, while similarities of Enneper's surface are *constant* curvature surfaces. This seems somehow dissatisfying considering the following simple connection between the geometry of a surface with that of its spin curve:

**Proposition 28.** *Let* $\mathbf{x}$ *be a minimal surface with spin curve* $\boldsymbol{\psi}$*. Then* $\boldsymbol{\psi}$ *is a line if and only if* $\mathbf{x}$ *is a plane.*

*Proof.* Let $z$ be a complex coordinate.

$$\begin{aligned}
\boldsymbol{\psi} \text{ is a line} \quad &\Leftrightarrow \quad \boldsymbol{\psi} = \mathbf{a}f(z) + \mathbf{b}, \quad f(z) \text{ holomorphic}, \ \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \ \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbf{C}^2 \\
&\Leftrightarrow \quad \boldsymbol{\psi}' = \mathbf{a}f'(z) = \begin{pmatrix} a_1 f'(z) \\ a_2 f'(z) \end{pmatrix} = \begin{pmatrix} p(z) \\ q(z) \end{pmatrix} \\
&\Leftrightarrow \quad g(z) = \frac{p(z)}{q(z)} = \frac{a_1}{a_2} \ \text{constant} \\
&\Leftrightarrow \quad \mathbf{x} \ \text{is a plane}
\end{aligned}$$

□

We can generalize this observation and simplify the expression for the curvature by repeating the frame analysis from a more natural point of view. Namely, we use a complex extension of the classical affine curve theory of Blaschke [1] et al, modifying the terminology to suit our purposes. Consider the $SL_2(\mathbf{C})$-invariant bilinear form $\| \cdot \|$ defined by

$$\begin{aligned} \|\cdot\| \,:\, C^\infty(M^2;\mathbf{C}^2) &\longrightarrow C^\infty(M^2;\mathbf{C}) \\ \boldsymbol{\psi} &\longmapsto det(\boldsymbol{\psi},\boldsymbol{\psi}') \end{aligned}$$

By the analysis above, the speed of the spin curve $\boldsymbol{\psi} : M^2 \to \mathbf{C}^2$ can be expressed with respect to a local complex coordinate $(U, z)$ as $\|\boldsymbol{\psi}'\|^{1/3} = (qp' - p\,q')^{1/3}$, where $p\,(z)$ and $q(z)$ are the component functions of $\boldsymbol{\psi}'$. The pseudo-arc parameter is given by $\sigma(z) = \int_{z_0}^{z} \|\boldsymbol{\psi}'\|^{1/3}\, dz$, and to say that $\boldsymbol{\psi}$ is parameterized by pseudo-arclength means that the curve has unit speed $\|\boldsymbol{\psi}'\| = det(\boldsymbol{\psi}', \boldsymbol{\psi}'') = 1$. Assume that this is the case, and consider the frame $F = (\boldsymbol{\psi}, \boldsymbol{\psi}', \boldsymbol{\psi}'')$. This is automatically third order

$$\omega \;=\; F^{-1}dF \;=\; \begin{pmatrix} 0 & 0 & 0 \\ \|\boldsymbol{\psi}'\| & 0 & \|\boldsymbol{\psi}''\| \\ 0 & \|\boldsymbol{\psi}'\| & 0 \end{pmatrix} d\sigma \;=\; \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & \|\boldsymbol{\psi}''\| \\ 0 & 1 & 0 \end{pmatrix} d\sigma$$

and shows that the curvature of a unit speed curve takes the form $\|\boldsymbol{\psi}''\| = q'p'' - p'q''$. Since the speed of a minimal curve is a power of the speed of its spin curve, unit speed curves correspond, and so the curvature $k$ of a minimal surface must also take this simple form when parameterized by pseudo-arclength.

Thus despite the appearance of terms that involve third order derivatives of $p\,(z)$, $q(z)$ in the expansion of Cartan's invariant $k$ and in that of the Schwartzian derivative $\mathcal{S}_z(g)$ of the Gauss map (also an invariant), there exists a parametrization with respect to which only *second* order terms in $p\,(z)$ and $q(z)$ appear. We might summarize this observation by saying that the Goursat transform is a second order effect on the level of spinors. Also, since $\|\boldsymbol{\psi}''\| = 0$ if and only if the tangent vectors $\boldsymbol{\psi}'' \,\|\, \boldsymbol{\psi}'''$ are parallel (i.e. linearly dependent), we see that $k$ is in fact a measure of the failure of $\boldsymbol{\psi}'$ to be a complex line. This is our "geometric interpretation" of Cartan's invariant. Motivated by this, we make the following definition:

**Definition 29.** Given a regular holomorphic curve $\boldsymbol{\psi} : M^2 \to \mathbf{C}^2$ and local complex coordinate $(U, z)$, the function $k_i : M^2 \to \mathbf{C}$ defined by $k_i(z) = \left\| \boldsymbol{\psi}^{(i)}(z) \right\|$ is called the $i^{th}$ *spin curvature* of $\boldsymbol{\psi}$ *with respect to the local coordinate* $(U, z)$, where $\boldsymbol{\psi}^{(i)}$ is the $i^{th}$ derivative.

These "spin curvatures" only make sense with respect to a fixed coordinate (since they are far from invariant under reparametrization) and are therefore not intrinsic to the image curve. But they do capture a rough notion of curvature, as pointed out above: $k_i(z) = 0$ iff $\boldsymbol{\psi}^{(i-1)}$ is a line. We will call the first two of these $k_1, k_2$ the *first and second spin curvatures* of a minimal curve (or surface), since by $SL_2(\mathbf{C})$-invariance, they are a complete set of invariants for the Goursat transform (two minimal curves (or surfaces) are Goursat equivalent if and only if their first and second spin curvatures agree with respect to a common complex coordinate). Moreover, they provide the desired generalization of proposition 13: the first spin curvature $k_1$ (which measures the failure of $\boldsymbol{\psi}$ to be a line[8]) is exactly the Hopf differential $II^{2,0}$ (which measures the failure of $\mathbf{x}$ to be a plane). Cartan's curvature does not capture this because it is the second spin curvature (when $z$ is the psuedo-arc parameter), which measures the failure of $\boldsymbol{\psi}$ to be a "quadratic line" (i.e. a curve of the form $\mathbf{a}f(z) + \mathbf{b}z + \mathbf{c}$ where $f'(0) = f(0) = 0$).

It is instructive to consider the constant curvature examples, which turn out to be familiar elementary affine varieties. Let $z$ be the arclength parameter for a curve $\boldsymbol{\psi}$ of constant non-zero curvature $k$. Then $F = (\boldsymbol{\psi}, A) = (\boldsymbol{\psi}, \boldsymbol{\psi}', \boldsymbol{\psi}'')$ is a third order frame such that

$$A^{-1}dA = \begin{pmatrix} 0 & k \\ 1 & 0 \end{pmatrix} = X$$

[8] Another way to interpret this fact is to observe that if $\boldsymbol{\psi}$ where a *real* planar curve, $\boldsymbol{\psi} : \mathbf{R} \to \mathbf{R}^2$, then $\|\boldsymbol{\psi}'\|$ is the numerator of the local expression for curvature, and the two agree in the unit speed case, which is a concrete example of how affine geometry generalizes Euclidean geometry.

But this equation is also satisfied by $A(z) = e^{Xz}$ since $X$ is a constant matrix. Notice that

$$X = \begin{pmatrix} 0 & k \\ 1 & 0 \end{pmatrix} = k^{1/2} \begin{pmatrix} 0 & k^{1/2} \\ k^{-1/2} & 0 \end{pmatrix}$$

$$\Rightarrow \quad X^{2j} = (k^{1/2})^{2j} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{and} \quad X^{2j+1} = (k^{1/2})^{2j+1} \begin{pmatrix} 0 & k^{1/2} \\ k^{-1/2} & 0 \end{pmatrix}$$

$$\text{so that} \quad (\boldsymbol{\psi}', \boldsymbol{\psi}'') = A = e^{Xz} = \sum_{j=0}^{\infty} \frac{X^{2j}\, z^{2j}}{(2j)!} + \sum_{j=0}^{\infty} \frac{X^{2j+1}\, z^{2j+1}}{(2j+1)!}$$

$$= \begin{pmatrix} \cos(k^{1/2}z) & 0 \\ 0 & \cos(k^{1/2}z) \end{pmatrix} + \begin{pmatrix} 0 & k^{1/2}\sin(k^{1/2}z) \\ k^{-1/2}\sin(k^{1/2}z) & 0 \end{pmatrix}$$

$$\Rightarrow \quad \boldsymbol{\psi}'(z) = \begin{pmatrix} \cos(k^{1/2}z) \\ k^{-1/2}\sin(k^{1/2}z) \end{pmatrix} \quad \Rightarrow \quad \boldsymbol{\psi}(z) = \begin{pmatrix} k^{-1/2}\sin(k^{1/2}z) \\ -k^{-1}\cos(k^{1/2}z) \end{pmatrix}$$

which parameterizes the affine quadric $k\, z_1^2 - k^2 z_2^2 = 0$, a complex hyperbola.

# Chapter 3

# The Goursat Transform for CMC1 Surfaces

## 3.1 Quaternionic Upper-Half Space

Before coming to our Goursat transform for CMC1 surfaces, let us introduce a model for hyperbolic 3-space in which this transform will be most easily understood. It seems an inevitable difficulty that distinguishes the study of space forms with non-negative curvature from those with negative curvature: the former each have a unique, natural, and easily visualized model, while the later have many natural models (and some rather exotic ones). A choice of model depends on which geometric phenomena one wishes to view, the various special features of hyperbolic geometry appearing more or less pronounced according to the model. There are two such features that particularly interest us at the moment. The first is the beautiful geometric interplay between hyperbolic space and its infinity boundary: any conformal transformation of the infinity boundary $\partial_\infty H^n \cong S^{n-1}$ extends to an isometry on all of $H^n$ (the "Poincaré extension"), and conversely an isometry of $H^n$ induces a conformal transformation on $S^{n-1}$. The second is that we already have a very satisfactory picture of the conformal boundary in the case $n = 3$: the action of $SL_2(\mathbf{C})$ on $S^2 = \widetilde{\mathbf{C}}$ by linear fractional transformation (LFT). We seek a model for $H^3$ whose infinity boundary is this specific conformal model of the 2-sphere, in an effort to obtain a more concrete picture of the way $SL_2(\mathbf{C})$ transformations act as hyperbolic isometries in dimension three.

As a first step, consider the upper-half plane. We have already discussed the action of $SL_2(\mathbf{C})$ on the extended complex plane, and it is not difficult to see that the subgroup which preserves the (extended) real line $\widetilde{\mathbf{R}} = \mathbf{R} \cup \{\infty\} \subset \mathbf{C} \cup \{\infty\} = \widetilde{\mathbf{C}}$ is exactly $SL_2(\mathbf{R})$. Of course, $\widetilde{\mathbf{R}} \cong S^1$, and the "circles" in $S^1$ are simply point pairs. Thus the Möbius transformations are nothing but the 1-1 maps - which certainly contains the $SL_2(\mathbf{R})$ transformations. One observes that if $Im\, z > 0$ then $Im(A \cdot z) > 0$ for any $A \in SL_2(\mathbf{R})$, so with Poicaré extension in mind, it is not hard to guess that the upper-half plane $U^2 := \{z \in \mathbf{C} \mid Im\, z > 0\}$ will be a model for $H^2$ with $SL_2(\mathbf{R})$ acting by isometries. Indeed, the familiar hyperbolic metric

$$ds^2 = \frac{dx^2 + dy^2}{y^2} = \frac{dzd\bar{z}}{y^2}$$

is immediately seen to be invariant under dilation $A = \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$, $A \cdot z = a^2 z$, translation $A = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$, $A \cdot z = z + b$, and inversion in the unit circle (followed by reflection through the y-axis) $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $A \cdot z = -\frac{\bar{z}}{|z|^2}$, which generate $SL_2(\mathbf{R})$:

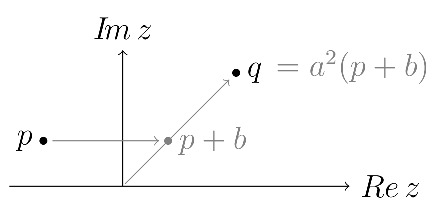


**Fig.3.1**

The first two transformations are sufficient to see transitivity (pictured in **Fig.3.1** above) and the isotropy subgroup at $z = i$ is $SO_2(\mathbf{R})$:

$$A \cdot i = \frac{ai+b}{ci+d} = i \quad \Leftrightarrow \quad ai + b = di - c \quad \Leftrightarrow \quad A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

so that $H^2 \cong SL_2(\mathbf{R})/SO_2(\mathbf{R})$. Dilation in this context is really "hyperbolic translation" in the y-direction, a notion that often seems bizarre and mysterious in other models, but which becomes totally transparent in upper-half space, as does the action of the remaining generators. Indeed, the upper-half space model allows us to easily imagine the correspondence between the respective transformations of $H^2$ and $\partial_\infty H^2$ by viewing the two as being attached, $H^2 \cup \partial_\infty H^2$, so that as one piece undergoes a transformation, the other moves in the most natural way to prevent tearing.

Let $\mathbf{H} = \{q = x_0 + x_1 i + x_2 j + x_3 k \mid x_i \in \mathbf{R}\}$ denote the space of quaternions, which forms a real non-commutative division algebra when the imaginary units $i, j, k$ satisfy the familiar relations $ij = k$, $jk = i$, $ki = j$, and $i^2 = j^2 = k^2 = -1$. Then we can consider $2 \times 2$ quaternionic matrices

$$M_2(\mathbf{H}) := \left\{ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \;\middle|\; a, b, c, d \in \mathbf{H} \right\}.$$

Conjugation of $q = x_0 + x_1 i + x_2 j + x_3 k$ is given by $\bar{q} = x_0 - x_1 i - x_2 j - x_3 k$, so the conjugate $\bar{A}$ and adjoint $A^* = \bar{A}^T$ of a quaternionic matrix are defined in the usual way, but non-commutativity prevents us from making basis-independent sense of the determinant. In the Hermitian case $A^* = A$ however, the order of multiplication in the determinant formula is irrelevant, which allowed Study to define a (real-valued) determinant of an arbitrary quaternionic matrix by $det_S(A) := det(A^*A)$. From this we can define $GL_2(\mathbf{H}) := \{A \in M_2(\mathbf{H}) \mid det_S(A) \neq 0\}$, which is exactly the set of $2 \times 2$ invertible matrices, or the set of matrices with $\mathbf{H}$-linearly independent rows and columns.

Similarly one defines the the subgroup $SL_2(\mathbf{H}) := \{A \in M_2(\mathbf{H}) \mid det_S(A) = 1\}$.

This group naturally acts on the compactification $\widetilde{\mathbf{H}} = \mathbf{H} \cup \{\infty\} \cong S^4$ by LFT:

$$A \cdot q \;=\; \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot q \;:=\; (aq+b)(cq+d)^{-1}$$

Obviously the subgroup preserving the (extended) complex plane $\widetilde{\mathbf{C}} \subset \widetilde{\mathbf{H}}$ is $SL_2(\mathbf{C})$, the Möbius transformations of $S^2$. In light of the upper-half plane model, it seems reasonable to hope that some upper-half 3-space contained in $\mathbf{H}$ will be preserved by this action, perhaps

$$U^3_j \;:=\; \{q = x_0 + x_1 i + x_2 j \;\mid\; x_2 > 0\}\,, \quad \text{or} \quad U^3_k \;:=\; \{q = x_0 + x_1 i + x_3 k \;\mid\; x_3 > 0\}$$

each of which, when given respective hyperbolic metrics

$$ds^2_j \;=\; \frac{dx_0^2 + dx_1^2 + dx_2^2}{x_2^2} \;=\; \frac{dq d\bar{q}}{x_2^2}, \qquad ds^2_k \;=\; \frac{dx_0^2 + dx_1^2 + dx_3^2}{x_3^2} \;=\; \frac{dq d\bar{q}}{x_3^2},$$

have the desired conformal model of $S^2$ as their infinity boundary $\partial_\infty U^3_j = \partial_\infty U^3_k = \widetilde{\mathbf{C}}$. In fact, *both* spaces are preserved by $SL_2(\mathbf{C})$, and more importantly, the action is transitive and isometric in each case. We will typically refer to $U^3_j$ as *the* quaternionic upper-half space, multiplication by $i$ providing an isometric "rotation" between the two $i : U^3_j \to U^3_k$.

**Lemma 30.** *$SL_2(\mathbf{C})$ acts transitively and isometrically by linear fractional transformation on the quaternionic half space $(U^3_j, ds^2_j)$ with isotropy subgroup $SU_2$ at $j$.*

*Proof.* This is a series of elementary computations analogous to those above for the upper-half plane, most of which follow from the commutation relation $zj = j\bar{z}$ for $z \in \mathbf{C}$, and the more general rule $\overline{q_1 q_2} = \bar{q}_2 \bar{q}_1$ for any $q_1$, $q_2 \in \mathbf{H}$. We represent an arbitrary element $q \in U^3_j$ by $q = z + xj$ where $z \in \mathbf{C}$ and $x \in \mathbf{R}^+$, and denote $A \cdot q = q_A = z_A + x_A j$. Then

$$\begin{aligned} q_A &= (aq+b)(cq+d)^{-1} \\ &= (aq+b)(\bar{q}\bar{c}+\bar{d})(\bar{q}\bar{c}+\bar{d})^{-1}(cq+d)^{-1} \\ &= \tfrac{a\bar{c}|q|^2+aq\bar{d}+b\bar{q}\bar{c}+b\bar{d}}{|cq+d|^2} \end{aligned} \qquad \begin{cases} aq\bar{d} \;=\; az\bar{d}+axj\bar{d} \;=\; a\bar{d}z+adxj \\ b\bar{q}\bar{c} \;=\; b\bar{z}\bar{c}-bxj\bar{c} = b\bar{c}\bar{z}-bcxj \end{cases}$$

$$\implies \quad q_A \;=\; \frac{a\bar{c}\,|q|^2+a\bar{d}z+b\bar{c}\bar{z}+b\bar{d}+xj}{|cq+d|^2} \quad \implies \quad \begin{cases} z_A = \frac{a\bar{c}|q|^2+a\bar{d}z+b\bar{c}\bar{z}+b\bar{d}}{|cq+d|^2} \\ x_A = \frac{x}{|cq+d|^2} \end{cases}$$

so that $x>0 \Rightarrow x_A>0$, and thus $SL_2(\mathbf{C})$ preserves $U_j^3$. Transitivity of this action is visually apparent, again using only horizonal translation $q = z+xj \;\mapsto\; q_A = (z+b)+xj$ and hyperbolic translation $q=z+xj \;\mapsto\; q_A = a^2(z+xj)$ with $a\in\mathbf{R}$, and the isotropy subgroup at $j$ is indeed $SU_2$ since:

$$A\cdot j = (aj+b)(cj+d)^{-1} = j \quad \Leftrightarrow \quad aj+b = j(cj+d) = -\bar{c}+\bar{d}j$$

$$\Leftrightarrow \quad c=-\bar{b}, \quad d=\bar{a} \quad \Leftrightarrow \quad A = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a}\end{pmatrix}, \quad |a|^2+|b|^2=1$$

Finally, the action is isometric since, writing $q_A = q_1q_2^{-1}$, where $\begin{cases} q_1 = aq+b \\ q_2 = cq+d \end{cases}$

$$\begin{aligned} dq_A &= dq_1q_2^{-1} - q_A dq_2 q_2^{-1} \\ &= [dq_1 - q_A dq_2]\,q_2^{-1} \\ &= [adq - q_A c\,dq]\,q_2^{-1} \\ &= [a-q_Ac]\,dq\,q_2^{-1} \end{aligned} \qquad \begin{aligned} a-q_Ac &= a-q_1q_2^{-1}c \\ &= a\,[c^{-1}q_2 - a^{-1}q_1]\,q_2^{-1}c \\ &= a\,[(q+c^{-1}d)-(q+a^{-1}b)]\,q_2^{-1}c \\ &= [ad-bc]\,c^{-1}q_2^{-1}c \;=\; c^{-1}q_2^{-1}c \end{aligned}$$

$$\Rightarrow \quad dq_A \;=\; c^{-1}q_2^{-1}c\,dq\,q_2^{-1} \quad \Rightarrow \quad A^*ds_j^2 \;=\; \frac{dq_Ad\bar{q}_A}{x_A^2} \;=\; \frac{\left(\frac{dqd\bar{q}}{|q_2|^4}\right)}{\left(\frac{x^2}{|q_2|^4}\right)} \;=\; \frac{dqd\bar{q}}{x^2} \;=\; ds_j^2.$$

□

Thus we again exhibit hyperbolic 3-space as the well-known homogeneous space $SL_2(\mathbf{C})/SU_2$. Again, it is not necessary to use the positive $j$-axis in this construction, nor even the usual complex plane $\mathbf{C}$. In general, one may choose any two orthogonal imaginary unit vectors $v, w \in S^2 \subset Im\,\mathbf{H}$, form the field $\mathbf{C}_v := span_{\mathbf{R}}\{1, v\}$, and let $SL_2(\mathbf{C}_v)$ act on the half space $U^3_w := \mathbf{C}_v \times \mathbf{R}^+ w$. Interestingly, there is a similar description for hyperbolic 5-space. Despite non-associativity, a completely analogous construction goes through in $\mathbf{O}$, the space of *octonions*, yielding an octonionic upper-half space $U^5_l = \mathbf{H} \times \mathbf{R}^+ l \simeq H^5 \simeq SL_2(\mathbf{H})/Sp_2$, where $l$ is the remaining (non-quaternionic) imaginary unit generator of $\mathbf{O}$, but we will have no use for this here.[1]

Standard isometries among the various models of hyperbolic 3-space help exhibit the naturality of the quaternionic model. We have already discussed the Hermitian matrix model

$$H^3_{Herm} = \{X \in M_2(\mathbf{C}) \mid X = X^*,\ detX = 1,\ trX > 0\} \subset Herm_2(\mathbf{C})$$

which is identified with the usual hyperboloid model $H^3_{H.loid}$ in Minkowski space after an isometric isomorphism of $Herm_2(\mathbf{C})$ with $\mathbf{R}^{1,3}$. The hyperboloid in turn stereographically projects to the Poincaré ball model $H^3_{P.ball}$ in $\mathbf{R}^3$ which, after a series of inversions, becomes the upper-half space $H^3_{U.half}$ (cf. Ratcliffe [18]). Collecting these isometries together, we have the explicit formulas in quaternionic notation:

[1]While indeed useful for the task at hand, it should perhaps be admitted that our model is strongly motivated out of deference to the quaternions in their own right, the author having been frequently struck by their geometric character and versatility. This sentiment is certainly nothing new: Hamilton himself famously spent years trying to make quaternions into a new paradigm for 3- and 4-dimensional Euclidean geometry, and these efforts (while never achieving their aim) have continued ever since by many mathematicians, in many contexts (cf. [19], [9], [5]). I feel that the natural presence of hyperbolic 2- and 3-spaces further underscores the impressively elegant setting that the quaternions seem to provide for low-dimensional space form geometry.

$$\begin{array}{rlcl} h_1 : & H^3_{Herm} & \longrightarrow & H^3_{H.loid} = \{\mathbf{x} \mid \langle \mathbf{x}, \mathbf{x} \rangle = -1,\, x_0 \geq 1\} \subset \mathbf{R}^{1,3} \simeq \mathbf{R} \times \mathbf{C} \times \mathbf{R} \\ & \begin{pmatrix} r & z \\ \bar{z} & s \end{pmatrix} & \longmapsto & \left( \frac{r+s}{2},\, z,\, \frac{r-s}{2} \right) \end{array}$$

$$\begin{array}{rlcl} h_2 : & H^3_{H.loid} & \longrightarrow & H^3_{P.ball} = \{\mathbf{x} \in \mathbf{R}^3 \mid \mathbf{x} \cdot \mathbf{x} =< 1\} \subset \mathbf{R}^3 \simeq \mathbf{C} \times \mathbf{R}j \\ & (t, z, x) & \longmapsto & \frac{z + xj}{1+t} \end{array}$$

$$\begin{array}{rlcl} h_3 : & H^3_{P.ball} & \longrightarrow & H^3_{U.half} = \{z + xj \mid x > 0\} \subset \mathbf{C} \times \mathbf{R}j \\ & (z, x) & \longmapsto & 2\frac{z+(1-x)j}{|z+(1-x)j|^2} - j \end{array}$$

the composition of which $h := h_3 \circ h_2 \circ h_1$ takes the simple form:

$$\begin{array}{rlcl} h : & H^3_{Herm} & \longrightarrow & H^3_{U.half} = U^3_j \\ & \begin{pmatrix} r & z \\ \bar{z} & s \end{pmatrix} & \longmapsto & \frac{z}{s} + \frac{1}{s}j \end{array}$$

Since each isometry above is equivariant with respect to the appropriate action by $SL_2(\mathbf{C})$ or $SO_{1,3}(\mathbf{R})$, $h$ must be as well: $h\left(AXA^*\right) = A \cdot h\left(X\right),\ \forall X \in H^3_{Herm},\, A \in SL_2(\mathbf{C})$. Otherwise said, the following diagram commutes:

$$\begin{tikzcd} SL_2(\mathbf{C}) \arrow[r, "id"] \arrow[d, "\pi_1"'] & SL_2(\mathbf{C}) \arrow[d, "\pi_2"] \\ H^3_{Herm} \arrow[r, "h", "\simeq"'] & U^3_j \end{tikzcd}$$

where $\begin{array}{rlcl} \pi_1 : & SL_2(\mathbf{C}) & \longrightarrow & H^3_{Herm} \\ & A & \longmapsto & AA^* \end{array}$ and $\begin{array}{rlcl} \pi_2 : & SL_2(\mathbf{C}) & \longrightarrow & U^3_j \\ & A & \longmapsto & A \cdot j \end{array}$

Note that $h(I) = j$, so we can compute $A \cdot j$ using either the formulas for $h$ and $AA^*$, or the general expression for $A \cdot q = q_A = z_A + x_A$ obtained in the proof of Lemma 1, setting $z = 0,\ x = 1$:

$$A \cdot j \ = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot j \ = \ \frac{(a\bar{c} + b\bar{d}) \ + \ j}{|c|^2 + |d|^2}$$

## 3.2 The transform

Note that there can be no non-trivial analogue of the Goursat transform for CMC1 surfaces based on our characterization from the previous chapter: If a CMC1 surface $f$ is specified by its hyperbolic Gauss map and Hopf differential $(g_h, II^{2,0})$,[2] our characterization would suggest that the transform surface $f_A$ should be that surface specified by $(A \cdot g_h,\ II^{2,0})$ where $A \in SL_2(\mathbf{C})$. However, the congruent surface $A \cdot f$ clearly *is* this transform, since $A : H^3 \to H^3$ acts as an isometry which takes a surface $f$ with Gauss map $g_h$ to a surface with Gauss map $A \cdot g_h$ and the same Hopf differential $II^{2,0}$. This phenomenon is a direct consequence of the fact that the Euclidean and hyperbolic Gauss maps have the same target space $\widetilde{\mathbf{C}}$, but only $SU_2$ Möbius transforms (of the unit sphere $\widetilde{\mathbf{C}} \simeq S^2 \subset \mathbf{R}^3$) extend to isometries on all of $\mathbf{R}^3$, while *any* $SL_2(\mathbf{C})$ Möbius transform (of the infinity boundary $\widetilde{\mathbf{C}} \simeq \partial_\infty H^3$) extends to an isometry of $H^3$.

Let us instead follow Goursat's reasoning to define a sensible analogue of the transform for CMC1 surfaces. Distasteful as it may be to proliferate jargon, we first introduce a bit of non-standard terminology to suggest the logical structure in which the transform is justified.

**Definition 31.** A *CMC1 curve* is a regular holomorphic map $F : M^2 \to SL_2(\mathbf{C})$ such that the pullback $F^*\omega = F^{-1}dF$ is null at each point $p \in M^2$ (i.e. $det(F^{-1}dF) \equiv 0$ on $M^2$).

[2] That a CMC1 surface can indeed be uniquely specified from the data $(g_h, II^{2,0})$ will be shown in section 3.4 on "the dual correspondence."

Another way to say this is that a CMC1 curve is a holomorphic map from $M^2$ to $SL_2(\mathbf{C})$ whose "Darboux derivative" takes values in the affine quadric $Q_2$ of non-zero null vectors

$$Q_2 := \{X \in \mathfrak{sl}_2(\mathbf{C}) - \{\mathbf{0}\} \mid det X = 0\}\,.$$

Of course, a CMC1 curve is then just a null holomorphic $SL_2(\mathbf{C})$-frame field along a CMC1 surface, specifically, along the surface defined by the projection $f = F \cdot j : M^2 \to H^3$. Conversely, by Bryant's result, any CMC1 immersion $f : M^2 \to H^3$ lifts to a CMC1 curve $F : M^2 \to SL_2(\mathbf{C})$ (the curve *associated to* $f$). Thus we have a correspondence between CMC1 surfaces and CMC1 curves (up to a constant left $SL_2(\mathbf{C})$-multiple).

Consider the adjoint action of $SL_2(\mathbf{C})$ on its Lie algebra $\mathfrak{sl}_2(\mathbf{C})$: $X \mapsto Ad_A(X) = AXA^{-1}$. Trace is invariant under conjugation (so that the action indeed preserves $\mathfrak{sl}_2(\mathbf{C})$), but so is the determinant (so that the action preserves the Killing form $\mathcal{K}(X) = -8\, det(X)$ on $\mathfrak{sl}_2(\mathbf{C})$), and thus this action preserves the quadric $Q_2$. Consequently, $A * F := FA^{-1}$ is a well-defined *left* action of $SL_2(\mathbf{C})$ on CMC1 curves, since if $F$ is such a curve, $det((A * F)^*\omega) = det(A(F^{-1}dF)A^{-1}) = det(F^{-1}dF) = det(F^*\omega) = 0$. This is our transform:

**Definition 32.** Let $f : M^2 \to H^3$ be a CMC1 surface and $F : M^2 \to SL_2(\mathbf{C})$ its associated CMC1 curve. A *Goursat transform* of $f$ is a CMC1 surface $f_A := FA^{-1} \cdot j$.

This definition indeed conforms to the original logic of the construction for minimal surfaces: given a surface, lift to the associated null curve, act by a linear transformation that preserves nullity, and project the result. Furthermore, in the same way that $\mathbf{x}_O$ is congruent to $\mathbf{x}$ when $O \in SO_3(\mathbf{R})$, the transform $f_A$ agrees with $f$ when $A \in SU_2$. As Goursat observed in the minimal case, we can isolate the "non-trivial part" of a fixed transform by factoring the matrix that accomplishes it, which for

CMC1 surfaces is simply the polar decomposition: if $A = UP$ with $U \in SU_2$ and $P$ positive definite, we have $f_A = FP^{-1} \cdot j$.

But the real justification for our definition is the following main result:

**Theorem 33.** *If the CMC1 surface $f$ corresponds to the minimal surface $\mathbf{x}$, then the transform $f_A$ corresponds to the transform $\mathbf{x}_{\Sigma(A)}$.*

*Proof.* Let $\boldsymbol{\gamma}$ be the minimal curve associated to $\mathbf{x}$. Consider the $\mathbf{C}$-linear map

$$\begin{array}{cccc} \Phi & : & \mathbf{C}^3 & \longrightarrow & \mathfrak{sl}_2(\mathbf{C}) \\ & & \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} & \longmapsto & \begin{pmatrix} v_3 & v_1 + iv_2 \\ v_1 - iv_2 & -v_3 \end{pmatrix} \end{array}$$

This is a vector space isomorphism, and the norm of a vector $\mathbf{v} \in \mathbf{C}^3$ with respect to the complexified Euclidean inner-product is given by $\|\mathbf{v}\| = -det(\Phi(\mathbf{v}))$. Thus $\Phi$ is conformal with respect to the bilinear form $\|\cdot\|$ on $\mathbf{C}^3$ and the Killing form $\mathcal{K}(\cdot)$ on $\mathfrak{sl}_2(\mathbf{C})$, and therefore restricts to a bijection of affine varieties $Q_1 \leftrightarrow Q_2$. In a sense, it is this map that *accomplishes* the minimal/CMC1 correspondence: to say that $f$ is the CMC1 cousin of $\mathbf{x}$ exactly means that the Darboux derivative of its associated curve $F$ equals the derivative of the associated curve $\boldsymbol{\gamma}$ composed with $\Phi$. In terms of the Weierstrass representation:

$$F^*\omega \;=\; F^{-1}dF \;=\; \begin{pmatrix} g & -g^2 \\ 1 & -g \end{pmatrix} \eta \;=\; \Phi \begin{pmatrix} \frac{1}{2}(1-g^2)\eta \\ \frac{i}{2}(1+g^2)\eta \\ g\eta \end{pmatrix} \;=\; \Phi(d\boldsymbol{\gamma})$$

*Claim*: $\Phi$ is a Lie algebra isomorphism satisfying $\Phi \circ \Sigma(A) = Ad_A \circ \Phi$ for $A \in SL_2(\mathbf{C})$.

*Proof of Claim*: When equipped with the usual cross product, $(\mathbf{R}^3, \times)$ is isomorphic as a Lie algebra to $(\mathfrak{so}_3(\mathbf{R}), [\,,\,])$ with the commutator bracket, via the map

$$\begin{array}{rcccc} \Theta_{\mathbf{R}} & : & (\mathbf{R}^3, \times) & \longrightarrow & (\mathfrak{so}_3(\mathbf{R}), [\ ,\ ]) \\ & & \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} & \longmapsto & \begin{pmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{pmatrix} \end{array}$$

This complexifies to an isomorphism $\Theta_{\mathbf{C}} : (\mathbf{C}^3, \times) \to (\mathfrak{so}_3(\mathbf{C}), [\ ,\ ])$, and a standard calculation shows that $\Theta_{\mathbf{C}}$ has equivariance $\Theta_{\mathbf{C}} \circ O = Ad_O \circ \Theta_{\mathbf{C}}$ for all $O \in SO_3(\mathbf{C})$. Define a new Lie bracket $\boxtimes$ on $\mathbf{C}^3$ by multiplying the cross product by a constant $\mathbf{v} \boxtimes \mathbf{w} := -2i(\mathbf{v} \times \mathbf{w})$. Then the algebra $(\mathbf{C}^3, \boxtimes)$ is still isomorphic to $(\mathfrak{so}_3(\mathbf{C}), [\ ,\ ])$

$$\begin{array}{rcccc} \Theta & : & (\mathbf{C}^3, \boxtimes) & \longrightarrow & (\mathfrak{so}_3(\mathbf{C}), [\ ,\ ]) \\ & & \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} & \longmapsto & \begin{pmatrix} 0 & 2iv_3 & -2iv_2 \\ -2iv_3 & 0 & 2iv_1 \\ 2iv_2 & -2iv_1 & 0 \end{pmatrix} \end{array}$$

and this isomorphism $\Theta := -2i\Theta_{\mathbf{C}}$ clearly also satisfies $\Theta \circ O = Ad_O \circ \Theta$. Next recall that the Lie algebra isomorphism $\Sigma_*$ induced by the 2:1 cover $\Sigma$ is given by

$$\begin{array}{rcccc} \Sigma_* & : & \mathfrak{sl}_2(\mathbf{C}) & \longrightarrow & \mathfrak{so}_3(\mathbf{C}) \\ & & \begin{pmatrix} x_3 & x_1 \\ x_2 & -x_3 \end{pmatrix} & \longmapsto & \begin{pmatrix} 0 & 2ix_3 & x_2 - x_1 \\ -2ix_3 & 0 & i(x_1 + x_2) \\ x_1 - x_2 & -i(x_1 + x_2) & 0 \end{pmatrix} \end{array}$$

and by virtue of the fact that $\Sigma$ is a group homomorphism, $\Sigma_* \circ Ad_A = Ad_{\Sigma(A)} \circ \Sigma_*$.[3] Then the inverse $\Sigma_*^{-1} : \mathfrak{so}_3(\mathbf{C}) \to \mathbf{C}^3$ is also a Lie algebra isomorphism satisfying $Ad_A \circ \Sigma_*^{-1} = \Sigma_*^{-1} \circ Ad_{\Sigma(A)}$. Setting $x_1 = v_1 + iv_2$, $x_2 = v_1 - iv_2$, we have $x_2 - x_1 = -2iv_2$, $i(x_1 + x_2) = 2iv_1$, showing that $\Phi$ is in fact the composition[4] $\Phi = \Sigma_*^{-1} \circ \Theta$. Thus we exhibit $\Phi$ as a Lie algebra isomorphism satisfying the desired equivariance.

[3] Given $X \in \mathfrak{g}$ and $\gamma : I \to G$ such that $\dot\gamma(0) = X$, we have $\Sigma(Ad_A\gamma) = \Sigma(A\gamma A^{-1}) = \Sigma(A)\Sigma(\gamma)\Sigma(A)^{-1} = Ad_{\Sigma(A)}\Sigma(\gamma)$, which differentiates at $t = 0$ to $\Sigma_*(Ad_A X) = Ad_{\Sigma(A)}\Sigma_*(X)$.

[4] Thus $\Phi$ "comes from" Goursat's 2:1 cover $\Sigma$, which, if we do regard $\Phi$ as somehow providing the minimal/CMC1 correspondence, strikes the author as an impressive coincidence considering the hundred year gap between the work of Bryant and that of Goursat!

$$\Phi \circ \Sigma(A) = \Sigma_*^{-1} \circ \Theta \circ \Sigma(A) = \Sigma_*^{-1} \circ Ad_{\Sigma(A)} \circ \Theta = Ad_A \circ \Sigma_*^{-1} \circ \Theta = Ad_A \circ \Phi$$

which proves the claim. $(\square)$

If a minimal surface $\mathbf{x} = Re\,\boldsymbol{\gamma}$ has CMC1 cousin $f = F \cdot j$, then $\mathbf{x}_{\Sigma(A)} = Re\,(\Sigma(A)\boldsymbol{\gamma})$ has cousin $\tilde{f} = \tilde{F} \cdot j$ where $\tilde{F}$ satisfies $\tilde{F}^*\omega = \Phi\,(\Sigma(A)d\boldsymbol{\gamma}) = A\Phi(d\boldsymbol{\gamma})A^{-1}$. But $(A * F)^*\omega = A\Phi(d\boldsymbol{\gamma})A^{-1} = \tilde{F}^*\omega$, so by Cartan-Darboux $\tilde{F} = A * F = FA^{-1}$ up to constant left-multiplication, and therefore $\tilde{f} = f_A$ up to rigid motion. $\square$

Regarding the Lawson correspondence as a map between (moduli) spaces of surfaces $\mathcal{L} : \mathcal{A} \to \mathcal{B}$ and the Goursat transform as an $SL_2(\mathbf{C})$ action on these spaces, the theorem says that $\mathcal{L}$ is equivariant with respect to this action. More generally, since this transform is defined in a totally analogous way for horospherical surfaces in any hyperbolic space, we say that it is an *equivariant deformation* of the horospherical surfaces with respect to the Lawson correspondence. This seems to make the correspondence at once more geometric and more perverse. More geometric in the sense that it respects the natural transformations of both sets of surfaces, more perverse in the sense that these transformations have quite different effects. For example, Gauss maps of corresponding surfaces rotate in "opposite directions" under the transform: $g \mapsto A \cdot g$ but $g_h = F \cdot g \mapsto FA^{-1} \cdot g$. Some of this phenomena will be evident in the following examples.

## 3.3 Examples

Beyond providing general insight into the structure of the Lawson correspondence, the main upshot of Theorem 4 is that it turns every computable example of a cousin pair $(\mathbf{x}, f)$ into a 3-parameter family of examples $(\mathbf{x}_{\Sigma(A)}, f_A)$. Unfortunately, only a few such examples are explicitly known, since only certain sets of Weierstrass data $(g, \eta)$ give rise to an ODE $F^{-1}dF = \tau$ which can be solved in closed form in terms of special functions. In an effort to expand our collection of computable pairs, we suggest a method for quickly obtaining these special data sets, after which the surfaces they generate can be graphed and studied.

Consider three holomorphic functions with no common zeros $r, s, t : M^2 \to \mathbf{C}$, and let $G : M^2 \to GL_2(\mathbf{C})$ be a solution of a (locally defined) differential equation of the form

$$G^{-1}dG = \begin{pmatrix} r & s \\ t & -r \end{pmatrix} dz \qquad \Longrightarrow \qquad \dot{G} = G \begin{pmatrix} r & s \\ t & -r \end{pmatrix}$$

We have already taken advantage of the special form of this first-order ODE system in the previous results, but let us now convert the system into a single *second-order* ODE, as first suggested in [20]. Let $(x, y)$ denote the first row of this matrix, so that the system becomes

$$\begin{pmatrix} \dot{x} & \dot{y} \\ * & * \end{pmatrix} = \begin{pmatrix} x & y \\ * & * \end{pmatrix} \begin{pmatrix} r & s \\ t & -r \end{pmatrix} \qquad \Longrightarrow \qquad \begin{cases} \dot{x} = rx + ty \\ \dot{y} = tx - ry \end{cases} \qquad \Longrightarrow \qquad \begin{cases} x = \frac{\dot{y}+ry}{s} \\ y = \frac{\dot{x}-ax}{t} \end{cases}$$

Differentiating the middle two equations and substituting for $\dot{x}$ and $\dot{y}$ as in the last two yields

$$\begin{cases} \ddot{x} - \left(\frac{\dot{t}}{t}\right) \dot{x} + \left(\frac{r\dot{t}-\dot{r}t}{t}\right) x = 0 \\[2ex] \ddot{y} - \left(\frac{\dot{s}}{s}\right) \dot{y} + \left(\frac{\dot{r}s-r\dot{s}}{s}\right) x = 0 \end{cases}$$

Then if $x_1, x_2$ and $y_1, y_2$ are pairs of linearly independent solutions of these two equations, respectively, then according to the above relations $G$ can be written

$$G = \begin{pmatrix} x_1 & \frac{\dot{x}_1 - a x_1}{t} \\ x_2 & \frac{\dot{x}_2 - a x_2}{t} \end{pmatrix} = \begin{pmatrix} \frac{\dot{y}_1 + r y_1}{s} & y_1 \\ \frac{\dot{y}_2 + r y_2}{s} & y_2 \end{pmatrix},$$

so to find $G$, it suffices to solve either of the two second order ODEs above. Since its Darboux derivative has zero trace, this map must have constant determinant, and thus $F = det(G)^{-\frac{1}{2}}\, G : M^2 \to SL_2(\mathbf{C})$ is a null holomorphic curve. Of course, we are interested in the spacial case $G^{-1}dG = \tau$ as in the Bryant construction. Given data $(g, \eta)$ on $M^2$, we abuse notation to locally express the 1-form as $\eta = \eta\, dz$, where $\eta$ now signifies a holomorphic function

$$\begin{pmatrix} r & s \\ t & -r \end{pmatrix} dz = \tau = \begin{pmatrix} g\eta & -g^2\eta \\ \eta & -g\eta \end{pmatrix} dz$$

in which case the second-order ODEs can be written

$$\begin{cases} \ddot{x} - (\frac{\dot{\eta}}{\eta})\, \dot{x} + (\dot{g}\eta)\, x = 0 \\ \ddot{y} - (\frac{g\dot{\eta} + 2\dot{g}\eta}{g\eta})\, \dot{y} + (\dot{g}\eta)\, x = 0 \end{cases}$$

We will usually only deal with the first of these two equations, but we refer to either equation as *the associated ODE* of the data $(g, \eta)$.

While local solutions are guaranteed to exist, these equations can get quite out of hand as soon as the data becomes the least bit complicated, so let us try to limit ourselves to some reasonable examples on the punctured sphere. To this end, $(g, \eta)$ will be called $(P, Q)$-*data* if the associated ODE takes the form $\ddot{x} + P\, \dot{x} + Q\, x = 0$ (or $\ddot{y} + P\, \dot{y} + Q\, y = 0$). Suppressing additive constants, a calculation shows that the possible sets of $(P, Q)$-data satisfy either[5]

[5]If the data and associated ODE are lifted to the universal cover $\pi : \widetilde{M}^2 \to M^2$ then the differential of the covering map plays a role and these formula for $(g, \eta)$ are not the most general, as we will see momentarily.

$$\begin{cases} g = \frac{\int Q\, e^{\int P}}{k} \\ \eta = -\frac{k}{e^{\int P}} \end{cases} \qquad \text{or} \qquad \begin{cases} g = \frac{1}{k\int Q\, e^{\int P}} \\ \eta = \frac{k(\int Q\, e^{\int P})^2}{e^{\int P}} \end{cases}$$

where $k \in \mathbf{C}$ is an arbitrary constant. Note that $Q = 0$ if and only if $g$ is constant (plane/horosphere), and $P = 0$ if and only if $\eta$ is constant. Besides these special cases, one might next look for data making $(P, Q)$ constant. Such $(P, Q)$-data are of the form $(g, \eta) = (P^{-1}e^{Pz}, -ke^{-Pz}dz)$ with $k = Q$. Making the change of variables $w = Pe^{-Pz}$ puts $(g, \eta) = (w^{-1}, \tilde{k}\, dw)$ where $\tilde{k} = Q/P^2$. This is exactly the data for associates of the Catenoid. Bryant dealt with this example in [4], using the characteristic polynomial to solve the associated ODE in terms of trigonometric functions.

*Example* 5. (Catenoid Cousins):

As Bryant observed, the associated ODE is sensitive to the constant $k$ used in the Catenoid data $(z^{-1}, k\, dz)$, so much so that the cousins of different conformal associates may appear radically dissimilar (for example, some catenoid cousins are surfaces of revolution, others are not). This is a general feature of the associated ODE and the correspondence itself, which is geometrically bizarre since, for example, *real* conformal transforms on the minimal side are nothing but homothety, the effect of which is so innocent that surfaces related in this way are implicitly regarded as equivalent in the literature. In the case of the Catenoid, we find that choosing $k$ small makes for a clearer picture:

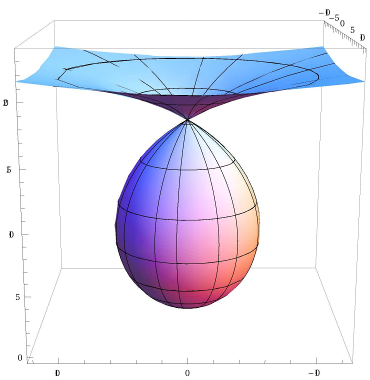

**Fig.3.2** $k = 0.1$

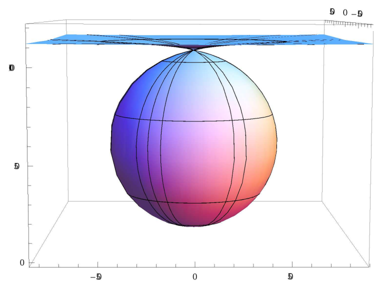

**Fig.3.3** $k = 0.01$

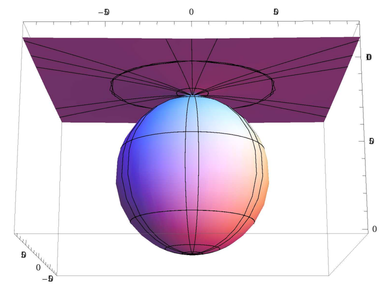

**Fig.3.4** $k = 0.01$

The two ends of the Catenoid correspond on the CMC1 cousin to ends which are asymptotic to horospheres. Here we see the bottom end, which is asymptotic to a horosphere with center $0 \in \widetilde{\mathbf{C}}$, attached by a very thin neck to the top end, which is also asymptotic to a horosphere with center $\infty \in \widetilde{\mathbf{C}}$ (which is a horizonal plane). Neck thickness is controlled by the parameter $k$, as it does for the neck of the Catenoid, as well as the asymptotic behavior of the ends. We can see that choosing $k$ small makes the bottom end appear very spherical and the top end very planar. Observe the effect of parabolic Goursat transforms on the $k = 0.01$ cousin:

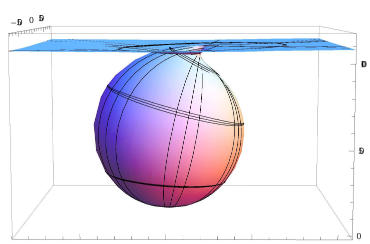

**Fig.3.5** $b = \frac{i}{4}$

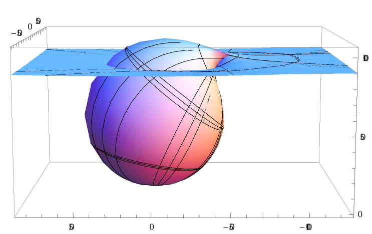

**Fig.3.6** $b = \frac{i}{2}$

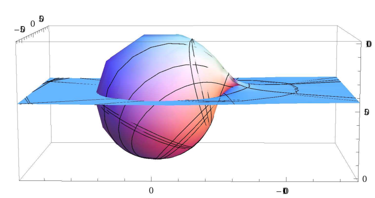

**Fig.3.7** $b = i$

Unlike in the minimal case, these transforms do not change the direction of the ends (the asymptotic limit on the infinity boundary), but they do change the *position* (the size and shape of the asymptotic horosphere), so that the ends move parallel to themselves under the transformation. The self-intersection and various deformed horospherical shapes are quite typical of CMC1 surfaces. Compare these images with the corresponding transforms of the Catenoid in the previous chapter.

Next consider the well-studied ODE $\ddot{x} - (\frac{2z}{1-z^2})\,\dot{x} + (\frac{k(k+1)}{1-z^2})\,x = 0$. This is *Legendre's equation*, solved by the Legendre functions, which are polynomials for integral $k \in \mathbb{Z}$. Setting $(P, Q) = (-\frac{2z}{1-z^2}, \frac{k(k+1)}{1-z^2})$ one can calculate from our formula above that $(g, \eta) = (z, -\frac{k(k+1)}{1-z^2}) = (z, \frac{k(k+1)}{(z+1)(z-1)})$ is $(P, Q)$-data on the twice-punctured place. This gives precisely what we are calling the Voss surface.

*Example* 6. (Voss Cousins):

Cousins of the $k = 1$ conformal associate have a crescent-shaped central piece

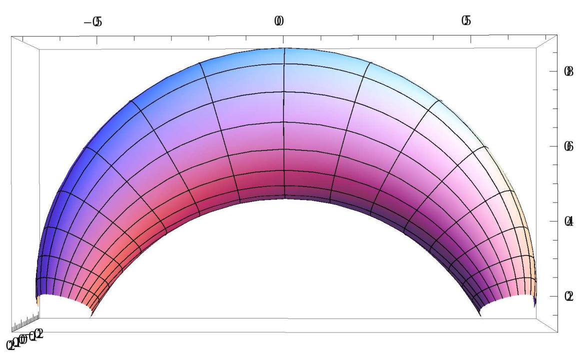

**Fig.3.8**

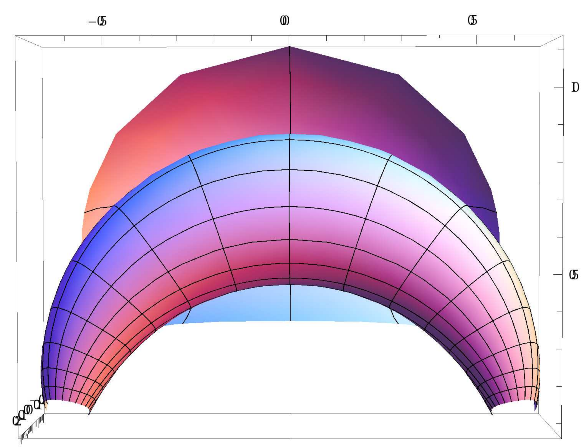

**Fig.3.9**

The surface has a vertical symmetry plane just like its minimal cousin, and the two downward-pointing ends here correspond to the two parallel upward-pointing ends we saw before on either side. These wrap around and intersect as they do on the cousin, forming the end at infinity, which appears as a portion of a horosphere at the back of the crescent.[6] Parabolic Goursat transforms break the vertical symmetry:

[6]These images also suggest that the Voss cousin possesses an extra horosphere of sphere of symmetry (which has no analogue on the Euclidean side) which would intersect the surface along the curve of self-intersection at the back of the crescent, but the author has not proven its existence analytically.

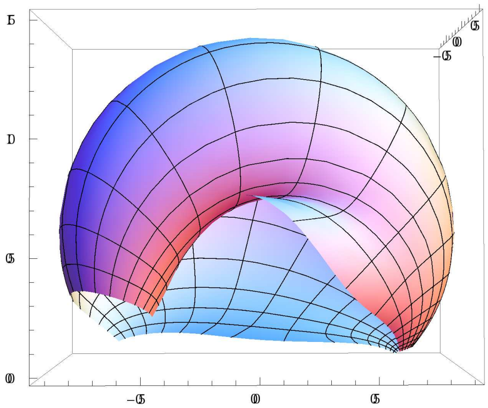

**Fig.3.10** $b = 1$

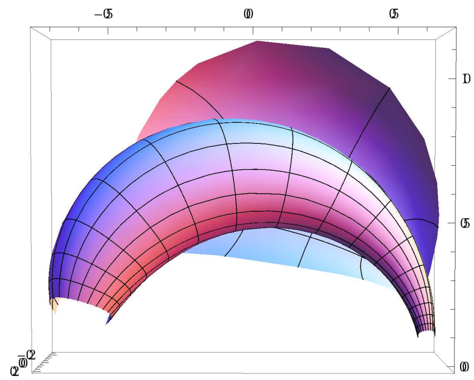

**Fig.3.11** $b = \frac{i}{2}$

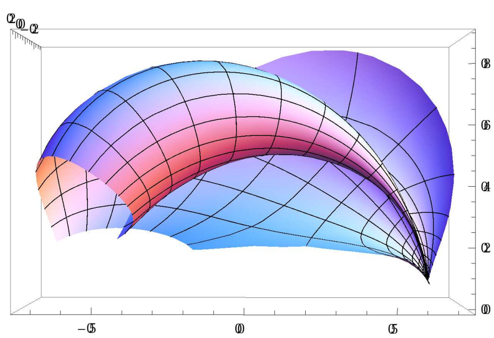

**Fig.3.12** $b = i$

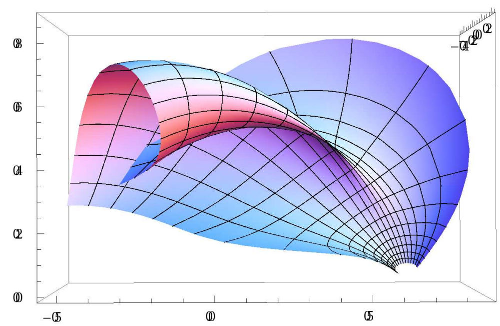

**Fig.3.13** $b = \frac{3i}{2}$

as they did in the minimal case, while transformation by diagonal elements preserve the symmetry too well to allow insightful observation, unlike in the minimal case. Again, the ends visibly change position under this deformation, but not direction - the asymptotic behavior remains fixed. Taking $b$ real makes the crescent thicken into a semi-horosphere, while $b$ imaginary unwraps it into the horosphere at the back.

The ODE $\ddot{x} - (\frac{1}{z})\,\dot{x} + x = 0$ is called *the zero-order Bessel equation*, which is the associated ODE for data $(g, \eta) = (\frac{z^2}{k}, \frac{k\,dz}{z})$, hence the name. It can be solved in closed form using Bessel's functions and spherical harmonics.

*Example* 7. (B-Cousins):

Like the Voss Cousins, the *Bessel Cousins* share several features with their minimal counterparts: the three catenoidal-necks of the minimal surfaces become three deformed horospheres here, two on the sides (corresponding to the upper necks) and one in the center (corresponding to the lower neck) only part of the interior of which can be seen.

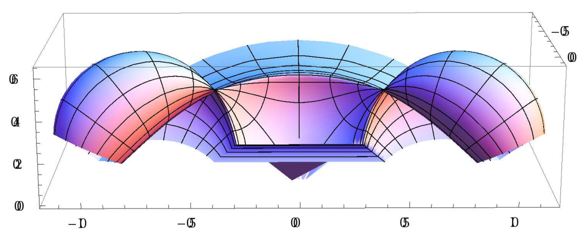

**Fig.3.14** front

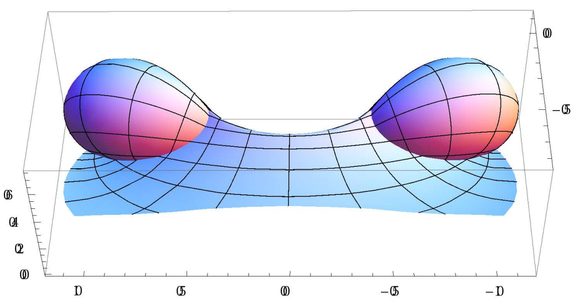

**Fig.3.15** back

The following parabolic transforms can be compared directly with those of the minimal cousins. Note particularly how the directions of rotation are opposite:

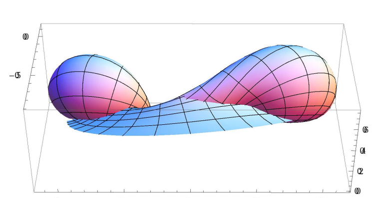

**Fig.3.16** $b = \frac{i}{2}$

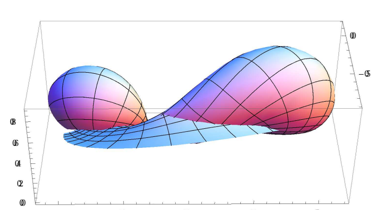

**Fig.3.17** $b = \frac{3i}{4}$

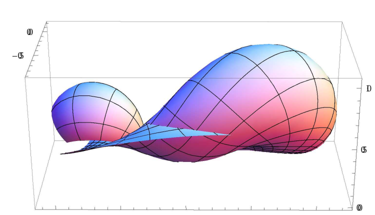

**Fig.3.18** $b = i$

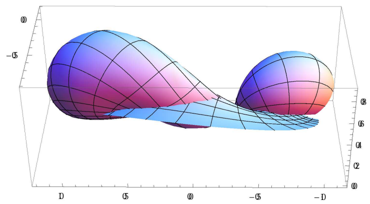

**Fig.3.19** $b = -\frac{3i}{4}$

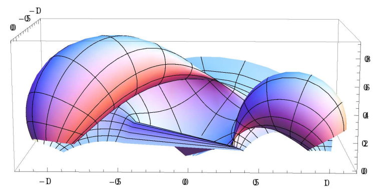

**Fig.3.20** $b = \frac{3i}{4}$

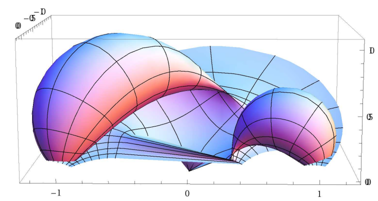

**Fig.3.21** $b = i$

The central piece rotates in the opposite direction when the inverse matrix is used (**Fig.3.19**), and frontal views exhibit the distortion of the central deformed horospherical part (**Fig.3.20** and **Fig.3.21**).

## 3.4 The dual correspondence

Theorem 4 underscores the fact that the Lawson correspondence is only a map between moduli spaces, not between sets of individual surfaces, which some readers might find troubling (if only for aesthetic reasons). However, there is an alternative correspondence between minimal and CMC1 surfaces (suggested in the opening discussion of section 3.2) which *can* be viewed as a literal map of surfaces. We call this the *dual correspondence* and argue, using Theorem 4, that it is the more natural correspondence in a certain sense.

**Definition 34.** Let $F : M^2 \to SL_2(\mathbf{C})$ be a CMC1 curve. The curve $F^{-1} : M^2 \to SL_2(\mathbf{C})$ is called the *dual* curve to $F$.

The dual curve is indeed CMC1, since its Maurer-Cartan derivative is $(F^{-1})^*\omega = -dFF^{-1}$, which has the same determinant as $F^*\omega = F^{-1}dF$. The terminology comes from Umehara and Yamada in [21], where given a CMC1 surface $f = F \cdot j$, they call $\hat{f} := F^{-1} \cdot j$ the *dual surface* to $f$. An equivalent way to regard this duality is the following: if a CMC1 curve $F$ satisfies the left-invariant ODE $F^*\omega = \tau$ with $\tau \in \bigwedge^{1,0}(M^2; Q_2)$, then its dual CMC1 curve $\tilde{F}$ satisfies the *right*-invariant ODE $\tilde{F}^*\tilde{\omega} = \tau$, where $\tilde{\omega}$ is the right-invariant Maurer-Cartan invariant 1-form (and vice versa), so that dual surfaces are just projections of these curves. One computational advantage of dual surfaces is that their period problems are easier to deal with, which helps facilitate construction of CMC1s from Weierstrass data (although the resulting surfaces need not be cousins of known minimal surfaces). But they also have the following intriguing geometric feature, which also first appeared in [20]:

**Proposition 35.** *If $f = F \cdot j$ has Weierstrass data $(g, \eta)$, then the dual surface $\hat{f} = F^{-1} \cdot j$ has hyperbolic Gauss map $\hat{g}_h = g$.*

*Proof.* Recall that $f = F \cdot j$ has hyperbolic Gauss map $g_h = [\dot{F}_1]$. Since $F = (F_1\, F_2) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ satisfies $\dot{F} = F\begin{pmatrix} g & -g^2 \\ 1 & -g \end{pmatrix}\eta$, we have $g_h = \frac{\dot{a}}{\dot{c}} = F \cdot g$. Then $\hat{g}_h = \frac{\dot{d}}{-\dot{c}} = \frac{-g(cg+d)\eta}{-(cg+d)\eta} = g$. $\square$

This fact[7] suggests the "dual correspondence:" a CMC1 surface $f$ is said to be the *dual cousin* of a minimal surface $\mathbf{x}$ if the hyperbolic Gauss map of $f$ agrees with the Gauss map of $\mathbf{x}$ and if they have the same Hopf differential. Thus the dual correspondence differs from the Lawson correspondence in that it replaces the use of the induced metric with the Gauss map. Both are based on criteria that completely determine the geometry of a surface, but the fundamental forms only determine a surface up to rigid motion. This is why the dual correspondence is arguably the more geometric of the two: it lifts from a bijection of moduli spaces to a bijection of

[7]This proposition shows that specifying the hyperbolic Gauss map $g_h$ and Hopf differential $II^{2,0}$ is sufficient to determine a CMC1 surface: if we use Weierstrass data $(g, \eta) = (g_h, \frac{II^{2,0}}{dg_h})$ to produce a CMC1 curve $F$ in the usual way, then $f = F^{-1} \cdot j$ is the (unique) CMC1 surface with hyperbolic Gauss map $g_h$ and Hopf differential $II^{2,0}$

surfaces. To see this, first observe as a trivial consequence of Theorem 4 (or by the first paragraph of section 3.2 together with Proposition 6):

**Proposition 36.** *Dual cousins of Goursat-equivalent minimal surfaces are congruent.*

*Proof.* Goursat-equivalent minimal surfaces $\mathbf{x}$ and $\mathbf{x}_{\Sigma(A)}$ have CMC1 cousins $f = F \cdot j$ and $f_A = FA^{-1} \cdot j$, which have duals $\hat{f} = F^{-1} \cdot j$ and $\hat{f}_A = (FA^{-1})^{-1} \cdot j = AF^{-1} \cdot j = A \cdot \hat{f}$. $\square$

In particular, if two minimal surfaces are congruent by a rotation $\Sigma(A) \in SO_3(\mathbf{R})$ of $\mathbf{R}^3$, where $A \in SU_2$, then their dual cousins are congruent by the "rotation" $A$ of $H^3$. This will allow us to promote the dual correspondence to a bijection between sets of surfaces. First, we provide initial value conditions that break the ambiguities introduced by the Weierstrass representation. We consider two sets of Weierstrass data $(p, q), (\tilde{p}, \tilde{q})$ on $M^2$ to be equivalent if they are related by a special linear transformation $A \in SL_2(\mathbf{C})$

$$\begin{pmatrix} p \\ q \end{pmatrix} \sim \begin{pmatrix} \tilde{p} \\ \tilde{q} \end{pmatrix} \quad \Leftrightarrow \quad \begin{pmatrix} p \\ q \end{pmatrix} = A \begin{pmatrix} \tilde{p} \\ \tilde{q} \end{pmatrix}$$

and fix a representative $[p, q]$ from each equivalence class. Next, fix a base point $z_0 \in M^2$ and insist that the minimal curve $\boldsymbol{\gamma}$ determined by $[p, q]$ satisfy $\boldsymbol{\gamma}(z_0) = \mathbf{0}$, and that its corresponding CMC1 curve $F$ satisfy $F(z_0) = I$. Then each data set representative gives rise to a fixed pair of dual cousins $(\mathbf{x}, f)$, all of whose congruent surfaces $(\Sigma(A)\mathbf{x},\ A \cdot f)$ are in fact dual cousins according to Proposition 7.

This suggests a new interpretation of Cartan's invariant (which we have renamed the *spin curvature*). Suppose we agree to construct CMC1 surfaces starting from spinor data $(p, q)$ and *right-invariant* curves (i.e. holomorphic solutions $F : M^2 \to SL_2(\mathbf{C})$ of ODE $dFF^{-1} = \tau$), which is the more convenient construction for the purposes of dealing with period problems. Then CMC1 surfaces with the same Hopf differential are congruent if and only if they have the same spin curvature. That is, spin curvature becomes a genuine notion of (metric) curvature in the CMC1 category, when the right-invariant representation is regarded as *the* construction method.

# Chapter 4

# Connections to Physics

## 4.1 Dirac's electron

We begin with a brief retelling of a very famous story from twentieth century physics: how the properties of "spin" were uncovered through an elegant combination of quantum mechanics and relativity. From a purely mathematical perspective, classical mechanics consists of a manifold $M$ ("configuration space" - the set of all possible states of a physical system) and an algebra $C^\infty(M)$ of real-valued functions on $M$ (the set of "observables" - the measurable physical quantities associated with the system). A distinguished function $H \in C^\infty(M)$ called the *Hamiltonian* then determines the evolution of the physical system over time (represented by a path $\gamma : \mathbf{R} \to M$) according to some geometrical prescription (Hamilton's equations). When Paul A.M. Dirac was still a graduate student in the 1920s, the appropriate mathematical formalism for quantum mechanics was still unclear, the "wave mechanics" of Schrodinger and "matrix mechanics" of Heisenberg being the primary contenders of the day. Dirac developed a formalism [7] that not only incorporated both of these approaches but also provided a sense in which the quantum picture "comes from" the classical one.

Mathematically, the general procedure is to replace the manifold of states $M$ with a Hilbert space $\mathbb{H}$, and the algebra of observables $C^\infty(M)$ with the algebra of Hermitian operators $Herm(\mathbb{H}) = \{A : \mathbb{H} \to \mathbb{H} \mid A = A^*\}$. The dynamics of the system (represented by a path $\psi : \mathbf{R} \to \mathbb{H}$ called the *wavefunction*) is again determined by one particular observable $H \in Herm(\mathbb{H})$ according to a geometric prescription (e.g. Schrödinger's equation). Many of the signature oddities of quantum phenomenology can then be interpreted as side effects of the fundamental setup: "quantization" of physical quantities, that is, discreteness of values assumed by observables, occurs because actual observed values are restricted to the spectrum of the corresponding operator; "non-commutativity" refers to the fact that multiplication in the algebra of quantum observables is no longer commutative;[1] the "probabilistic" character of the theory occurs when $\mathbb{H}$ is taken to be $L^2(M; \mathbf{C})$, so that the (normalized) squared-modulus of the wavefunction at a fixed time $|\psi_{t_0}|^2$ can be interpreted as a probability density function on space.

For simple physical systems, this formalism provides a very satisfying correspondence between classical and quantum observables, which makes essential use of an important geometric feature of classical mechanics: Nöther's theorem states that any *conserved* observable is associated with a *symmetry* (1-parameter group of transformations) of the physical system, and conversely every symmetry gives rise to a conserved quantity. In the language of physicists, a conserved quantity is associated with an infinitesimal generator (a vector in a Lie algebra, which generates the 1-parameter group of transformations via the exponential map). The classical/quantum correspondence of observables is then simply the identification of the observable with that

[1]That is, $C^\infty(M)$ with multiplication of functions defined point-wise is commutative, while $Herm(\mathbb{H})$ with multiplication of operators defined by composition is highly non-commutative. Of course, as *Lie algebras*, these are both (by definition) anti-commutative (where $C^\infty(M)$ is equipped with the Poisson bracket and $Herm(\mathbb{H})$ the commutator bracket), which is the fundamental *similarity* that first inspired Dirac.

generator (provided that this generator can be somehow realized as a Hermitian operator on some appropriate Hilbert space).[2] The fundamental examples are linear and angular momentum, which generate linear translations and rotations, respectively, as we expect on the basis of the momentum conservation laws.

Consider a system consisting of a single particle in $(\mathbf{R}^3, x^i)$. The coordinates of the particle's position $x^i$ and momentum $p_i$ have canonical[3] quantizations

$$x^i \;\leftrightarrow\; M_{x^i} \qquad\qquad p_i \;\leftrightarrow\; i\hbar\,\partial_i$$

where $M_{x^i}$ is the multiplication operator $M_{x^i} f(\mathbf{x}) = x^i f(\mathbf{x})$, and $\partial_i$ is the partial derivative operator $\partial_i f = \frac{\partial f}{\partial x^i}$. On the basis of energy conservation, we identify the total energy $E$ with the infinitesimal generator of time translation (as one does on the classical side), obtaining the correspondence $E \leftrightarrow i\hbar\,\partial_t$. On the other hand, in the absence of potential energy, the particle's total energy is kinetic, $K = \frac{1}{2}\,mv^2 = \frac{p^2}{2m}$, so that $E = K$ can *also* be quantized as $-\frac{\hbar^2}{2m}\,(\partial_1^2 + \partial_2^2 + \partial_3^2)$. Applying this operator to the particle's wavefunction $\psi$, we obtain the *Schrödinger equation*:

$$i\hbar\,\frac{\partial}{\partial t}\psi = -\frac{\hbar^2}{2m}\nabla^2\psi$$

which can be interpreted as the equation of motion for a quantum particle.

[2]A few of the implicit identifications and mathematical troubles common in this formalism bare mention: The Hilbert space $\mathbb{H}$ and its projectivization $\mathcal{P}\mathbb{H}$ are often used interchangably, since only normalized wavefunctions are considered physically relevant; Hermiticity and anti-Hermiticity are often used interchangably (multiplication by $i$ providing the identification), depending on whether one wishes to emphasize the operator's role as an observable (Hermitian operators possess a *real* spectrum) or its role as an infinitesimal generator (anti-Hermitian operators generate *unitary* transformations); The eigenvectors of relevant observables may only be *generalized* functions, since these operators often turn out to be unbounded or undefined on parts of $\mathbb{H}$. Hopefully the sympathetic reader will agree that the conceptual beauty of this formalism justifies the rather extreme efforts needed to maintain its mathematical consistency.

[3]This is also sometimes called "naive" quantization. While Nöther's theorem always plays an important role, the precise rules governing the correspondence of observables for more complicated systems (that is, more general manifolds $M$) are unclear or unknown. Finding procedures that provide the correspondence unambiguously in the general case is a subject of great interest and difficulty called *geometric quantization*.

Spin is an unusual quantum observable because it does not seem to have a precise classical analogue. Despite this, its theory is strongly motivated by such an analogy: we imagine our particle to be a tiny spinning sphere of positive radius (instead of a point particle) with charge $q$. This rotation will induce a magnetic moment $\mu = \frac{gq}{2m}S$, where $S$ is a vector pointing along the axis of rotation called the *spin vector*, and $g$ is some constant called the *gyromagnetic ratio*. This intuitive picture of spin turns out to be fundamentally inaccurate, but it assisted Pauli in ascertaining the correct mathematical properties: By regarding spin as the quantization of angular momentum, we may expect (following Nöther) that since the group of rotations on $\mathbf{R}^3$ is generated by the Lie algebra $\mathfrak{so}_3(\mathbf{R})$, the Hermitian operators for spin should form a *representation* of this algebra on some Hilbert space $\mathbb{H}$. The spin matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

constitute such a representation on $\mathbb{H} = \mathbf{C}^2$ (due to the fact that $i\sigma_j$ generate $\mathfrak{su}_2 \simeq \mathfrak{so}_3(\mathbf{R})$). Pauli therefore postulated that, for a spin-1/2 fermion (for example, an electron), the operator corresponding to the spin vector pointing in the $\mathbf{x} = (x_1, x_2, x_3) \in S^2$ direction is $S_{\mathbf{x}} = x^i \sigma_i$.

The first consequence of this postulate is that the (spin) wavefunction $\psi$ must now take values in $\mathbf{C}^2$, since[4] the Pauli matrices $\sigma_i$ are $2 \times 2$. Also recall the isometric isomorphism $S : \mathbf{R}^{1,3} \to Herm_2(\mathbf{C})$ called the Paul map in the introduction, and note that the assignment $\mathbf{x} \to S_{\mathbf{x}}$ is simply a restriction of this map. Then if $\Sigma : SU_2 \to SO_3(\mathbf{R})$ the usual double covering, we have seen that $S$ satisfies the equivariance law $S_{\Sigma(A)\mathbf{x}} = AS_{\mathbf{x}}A^*$. This is essentially what is needed to explain the famous and bizarre behavior that spin-1/2 particles exhibit: continuous rotation of

[4]This is the general case for 2-state systems. Spin-1/2 is such a system in the sense that, in any given direction, the spin vector can only point "up" (1-eigenvector) or "down" ($-1$-eigenvector). This characteristic quantum fact was the actual motivation for Pauli's investigation, contrary to the implication of our exposition.

the particle through angle of $2\pi$ has the effect of transforming the wavefunction to its negative, $\psi \to -\psi$.

The key to this behavior is the topology of the rotation group: Pauli's algebra representation of $\mathfrak{so}_3(\mathbf{R})$ on $\mathbf{C}^2$ does *not* integrate to a group representation of $SO_3(\mathbf{R})$, but to only to that of its double cover $SU_2$. That this representation of $SU_2$ does not descend to $SO_3(\mathbf{R})$ is what physicists mean by statements like "$\psi$ is a double-valued representation of $SO_3(\mathbf{R})$," or "$\psi$ is a *spinor*,[5] not a vector." Rather than repeat Pauli's treatment, we would like to suggest a more geometric demonstration that $\psi$ is a spinor: To say that the spin vector of an electron has direction $\mathbf{x} = (x^1, x^2, x^3) \in S^2$ means that $\psi$ is a 1-eigenvector of $S_{\mathbf{x}}$. One can check that the 1-eigenvector for $S_{\mathbf{x}}$ is $\psi = \begin{pmatrix} x^1 + ix^2 \\ 1 - x^3 \end{pmatrix}$, which can be normalized by identifying it with its complex span, or equivalently, with a point in projectivized Hilbert space, $[\psi] \in \mathbf{C}P^1$. We thus obtain a map from $S^2$ (the directions of the spin vector) to $\mathbf{C}P^1$ (the (projectivized) directions of the 1-eigenvectors of the associated operator), and this map is none other than stereographic projection (with the usual identification $\widetilde{\mathbf{C}} \leftrightarrow \mathbf{C}P^1$):

$$\begin{array}{rccc} \sigma \ : & S^2 & \longrightarrow & \mathbf{C}P^1 \\ & \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} & \longmapsto & \begin{bmatrix} \frac{x^1+ix^2}{1-x^3} \\ 1 \end{bmatrix} \end{array}$$

Regarding the wavefunction $\psi : \mathbf{R}^3 \to \mathbf{C}^2$ as a lift of the radial extension of $\sigma$, we obtain the famous equivariance condition $\psi(\Sigma(U)\mathbf{x}) = U\psi(\mathbf{x})$ (making the wavefunction a spinor) as a result[6] of the identical equivariance satisfied by $\sigma$. This includes as a simple consequence the sign-flip behavior discussed above: the path $\gamma : [0, \pi] \to SU_2$

[5]In physics, spinors are defined as those objects on which an orthogonal group acts via a double representation. The term is generalized in mathematics to mean those objects on which a Clifford algebra acts (cf. Lawson [14]).

[6]On the basis of this observation, it is tempting to conclude that, at least in the two-dimensional case, stereographic projection $\sigma$ *occurs* in nature! This also provides an ultimately physical justification for regarding $\mathbf{C}P^1$ as the "spin model" of the 2-sphere, in the sense of section 2.5.

defined by $\gamma(\theta) = diag(e^{i\theta}, e^{-i\theta})$ connects $\gamma(0) = I$ to $\gamma(\pi) = -I$ in $SU_2$, while $\Sigma \circ \gamma$ is a *closed* path in $SO_3(\mathbf{R})$. Applying $\Sigma \circ \gamma$ to $\mathbf{R}^3$, we obtain the result as $\theta$ varies from 0 to $\pi$: the particle rotates about the $z$-axis through a full $2\pi$ angle, while $\psi = \gamma(0)\psi$ changes into $-\psi = \gamma(\pi)\psi$ by the equivariance condition.

Let us return to the general setup for a moment. We obtained the Schrödinger equation by quantizing the classical energy relation $E = \frac{p^2}{2m}$. However, to incorporate relativity, this relation must be modified to account for the contribution of the particle's rest mass $m$: $E^2 = c^2p^2 + c^2m^4$. Letting $x^0 = t$ be the time coordinate in Minkowski space $\mathbf{R}^{1,3}$ and choosing units so that $c = 1$, the relativistic energy relation quantizes to the *Klein-Gordon equation*:

$$(\Box + M^2)\psi = 0$$

where $M = \frac{m}{\hbar}$ and $\Box = \partial_0^2 - \partial_1^2 - \partial_2^2 - \partial_3^2$ is the Laplacian in Minkowski space. This equation seems to have been first noticed by Schrödinger himself and was well-established by the time Dirac was considering the relativistic theory of the electron in 1927. Apparently, Dirac was troubled by the appearance of the second-order time derivative $\partial_0^2$ in the Klein-Gordon equation (for reasoning related to the probabilistic interpretation of the wavefunction $\psi$) and preferred the first-order time derivative $\partial_t$ as it appears in Schrödinger's equation. But a Lorentz invariant equation of motion which is first-order in one coordinate would have to be first-order in *all* coordinates, since there can be no preferred direction in space-time. Dirac therefore sought a first-order equation of motion for the electron that would still capture the physics of the Klein-Gordon equation. His solution involved a "square-root of the Laplacian," which leads to a surprising elaboration of Pauli's work on the mathematics of spin.

Define a modified version $\widetilde{S}$ of the Pauli spin map $S$ by

$$\begin{aligned} \widetilde{S} \; : \; \mathbf{R}^{1,3} &\longrightarrow Herm_2(\mathbf{C}) \\ x^i e_i &\longmapsto x^0\sigma_0 - x^i\sigma_i \end{aligned}$$

Whereas $S$ satisfied $S_{\Sigma(A)\mathbf{x}} = AS_{\mathbf{x}}A^*$ (extending $\Sigma$ to the 2:1 cover $\Sigma : SL_2(\mathbf{C}) \to SO_{1,3}(\mathbf{R})$ of the Lorentz group), $\widetilde{S}$ satisfies $\widetilde{S}_{\Sigma(A)\mathbf{x}} = A^{*-1}\widetilde{S}_{\mathbf{x}}A^{-1}$, as well as the property that $S_{\mathbf{x}}\widetilde{S}_{\mathbf{x}} = \widetilde{S}_{\mathbf{x}}S_{\mathbf{x}} = \|\,\mathbf{x}\|\,I$, where $\|\,\mathbf{x}\| = det\,S_{\mathbf{x}}$ is the Lorentzian norm. We can piece these together into a linear map

$$\begin{aligned} \gamma \;:\; \mathbf{R}^{1,3} &\longrightarrow \quad M_4(\mathbf{C}) \\ \mathbf{x} \quad &\longmapsto \begin{pmatrix} 0 & S(\mathbf{x}) \\ \widetilde{S}(\mathbf{x}) & 0 \end{pmatrix} \end{aligned}$$

which defines the *Dirac matrices* according to $\gamma^i := \gamma(e_i)$. Dirac engineered these matrices[7] specifically for their commutation relations $\gamma^i\gamma^j + \gamma^j\gamma^i = 2\eta_{ij}$, where $\eta_{ij} = diag(1,-1,-1,-1)$ (otherwise said, $\gamma$ identifies space-time $\mathbf{R}^{1,3}$ with (generators of) its own complexified Clifford algebra $Cl(\mathbf{R}^{1,3})\otimes\mathbf{C}$). He then defined the famous *Dirac operator* $\partial\!\!\!/ := \gamma(\nabla) = \gamma^i\partial_i$, a first order differential operator whose square is indeed the wave-operator $\partial\!\!\!/^2 = \Box$, on account of the (anti)commutativity of the $\gamma^i$. The *Dirac equation*

$$(\partial\!\!\!/ + iM)\psi = 0$$

is then the desired first-order equation of motion for the wavefunction of the electron, from which the Klein-Gordon equation follows by applying $(\partial\!\!\!/ - iM)$ to both sides.

An immediate consequence of this equation[8] is that the (spin) wavefunction $\psi$ must now take values in $\mathbf{C}^4$, since the Dirac matrices $\gamma^i$ are $4 \times 4$. Following Pauli, we may ask: how does this wavefunction transform under rotation? In the relativistic setting, rotation means Lorentz transformation, and we note that the equivariance laws for the maps $S$ and $\widetilde{S}$ imply that the map $\gamma$ satisfies $\gamma(\Sigma(A)\mathbf{x}) = [A]\,\gamma(\mathbf{x})\,[A]^{-1}$, where $[A] \in SL_4(\mathbf{C})$ is defined by

---

[7]These are also called the *gamma matrices* in the physics literature. We present them here in the *Weyl representation*, but there are several others.

[8]Other implications include an impressively accurate calculation of the gyromagnetic ratio $g$ and the theoretical anticipation of antimatter, just to name a few!

$$[A] := A \oplus A^{*-1} = \begin{pmatrix} A & 0 \\ 0 & A^{*-1} \end{pmatrix}$$

Since the Dirac operator can be written $\gamma(\nabla)$, it transforms in the same way under Lorentz transformation $\Sigma(A)$: $\partial\!\!\!/ \mapsto [A]\,\partial\!\!\!/\,[A]^{-1}$. The demand that the Dirac equation be Lorentz invariant then implies that the wavefunction $\psi$ should obey the transformation law $\psi(\Sigma(A)\mathbf{x}) = [A]\psi(\mathbf{x})$, making it a 4-spinor. The the form of $[A]$ prompted Dirac to write $\psi$ a direct sum of two 2-spinors $\psi_L, \psi_R \in \mathbf{C}^2$ which transform in with opposite "chirality":

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \quad \text{where} \quad \begin{cases} \psi_L(\Sigma(A)\mathbf{x}) = A\,\psi_L(\mathbf{x}) \\ \psi_R(\Sigma(A)\mathbf{x}) = A^{*-1}\,\psi_R(\mathbf{x}) \end{cases}$$

The Dirac equation can then be regarded as a coupled system of PDE in the left- and right-handed 2-spinors $\psi_L, \psi_R$:

$$\begin{cases} S(\nabla)\psi_R + iM\psi_L = 0 \\ \widetilde{S}(\nabla)\psi_L + iM\psi_R = 0 \end{cases}$$

Notice that these equation *de*couple when the mass is zero $M = 0$. Conjugating either equation by $J$ shows that these are actually the same equation, sometimes called the *Weyl equation*, which has been useful in the study of the neutrino (also a 1/2-spin particle, whose mass is extremely small, if not zero). We might summarize this result by saying that: the rotational equivariance of Pauli's 2-spinor wavefunction extends to a Lorentzian equivariance of Dirac's coupled left and right 2-spinor wavefunctions in two different ways.[9]

[9] These inequivalent "left" and "right" representations $\psi \to A\psi$ and $\psi \to A^{*-1}\psi$ of $SL_2(\mathbf{C})$ are irreducible, and other irreducible representations of $SL_2(\mathbf{C})$ can be built from tensor products of these two (cf. Bleecker [2]). The spin quantum number $s$ can then be interpreted mathematically as classifying elementary particles according to which representation of $SL_2(\mathbf{C})$ manifests in the Lorentz transformation of its wavefunction. Fermions (the "building blocks" of matter) have half-integral spin (like the $s = \frac{1}{2}$ electron) while bosons (the "glue" holding matter together) have integral spin (like the $s = 1$ photon).

## 4.2 The analogy with horospherical surfaces

We would like to reinterpret what has been done for horospherical surfaces in the previous chapters as analogous to what Dirac did for the electron, on the basis of which we suggest that certain surfaces might actually be thought of as particles. This will be of a completely hypothetical and rather vague nature, and will not involve any new mathematical results; the object is simply to summarize our work up to this point in terms of a physical picture, draw attention to the very special characteristics of horospherical surfaces, and collect evidence to suggest why they might find application in particle physics.

Identifying geometric objects with physical ones is of course nothing new. For example, in Yang-Mills theory, if the internal symmetries of a particle represent a group $G$, then phase space for the "internal motion" is a principal $G$-bundle over space (e.g. $G = SU_2$ for the spin of the electron, $G = SU_3$ for the isospin of the nucleon, etc). A connection in this bundle is then thought of as the potential for a physical field, and its curvature as the strength of that field. Running this analogy in reverse leads to treating certain principal bundles as particles (e.g. magnetic monopoles, instantons, etc) and their characteristic classes as physical properties (hence terminology like *topological charge*).

So in what sense are surfaces like particles? In general relativity, the motion of a point particle is recorded by its world-line (the path it sweeps out in space-time), and if the particle is not subject to external forces, this path is a geodesic with respect to the metric induced by the presence of matter/energy (stress-energy tensor). In the highly speculative string theory currently under development, a particle is represented by a vibrating string, so now the motion is recorded by its world-*sheet* (the surface its string sweeps out in space-time). If the particle is not subject to external forces, this surface ought to be minimal with respect to the induced metric. In fact, one of the many surprising and ambitious fundamental objectives of some versions of string theory is to obtain a model in which all the physics of a particle can be expressed through geometry of its world-sheet. From our perspective, the geometry of a surface is determined by the Gauss and Codazzi equations, i.e. its integrability conditions (in the sense of Bonnet's theorem). Thus if the string philosophy is to be taken seriously, it suggests that we think in terms of a correspondence like

$$\begin{array}{ccc} \{\mathit{Surfaces}\} & \longleftrightarrow & \{\mathit{Particles}\} \\ \\ \{\mathit{Integrability\ \ Conditions}\} & \longleftrightarrow & \{\mathit{Equations\ \ of\ \ Motion}\} \end{array}$$

where particles are identified with their world-sheets, and the equations that govern motion with the equations that govern geometric invariants. If a particle modeled on a *closed* string is created and annihilated, its world-sheet is compact, and if it is modeled on an infinite *open* string, the result is non-compact. In both cases, the topology of the surface is thought to play an important role in describing particle creation, annihilation, and interaction.

One of the immediate difficulties with making this picture precise is that the equations describing these objects can assume a variety of forms, depending on how they are modeled. On the physics side, we have already seen one of the most important examples: the Klein-Gordon equation and the Dirac equation are both valid forms of the Schrödinger equation, but each requires its own separate interpretation. Similarly, in surface theory, the structure equations can be expressed directly in terms of first order frames and the fundamental forms, or equivalently in terms of the induced complex structure and its associated invariants. Making the above correspondence consistent should depend critically on finding a coherent way to compare the relevant equations simultaneously.

The spinor representation seems to provide just such a device. Although we have not described it in full generality, it turns out that *any* conformal immersion $\mathbf{x} : M^2 \to \mathbf{R}^3$ can be represented in terms of a pair of (not necessarily holomorphic) spinor fields $(p, q)$ according to the same recipe we used for minimal surfaces

$$\mathbf{x} \;=\; Re \int \begin{pmatrix} \frac{1}{2}(q^2 - p^2) \\ \frac{i}{2}(q^2 + p^2) \\ qp \end{pmatrix}$$

and the structure equations can then be translated into conditions on these spinors. We have suggestively denoted the associated spin curve $\psi = \begin{pmatrix} p \\ q \end{pmatrix} : M^2 \to \mathbf{C}^2$ which, for the purposes of our analogy, we will now refer to as the *wavefunction* for the surface $M^2$. One may ask, as Pauli did, when the surface is rotated $\mathbf{x} \mapsto R\mathbf{x}$ by $R \in SO_3(\mathbf{R})$, how does its wavefunction $\psi$ transform? The answer is (by Lemma 2 of section 2.2, writing $R = \Sigma(U)$ with $U \in SU_2$) according to $\psi \mapsto U\psi$ (that is, exactly as a Pauli 2-spinor). Geometrically, this is because the Gauss map $g = [\psi]$ is the stereographic projection of the normal direction to the surface. But as we have just seen, the direction of the spin vector has exactly the same interpretation! This suggests, taking the surface/particle analogy to a rather literal extreme, that the normal directions of a surface in Euclidean can be thought of as the spin vector directions of the particle it represents (and conversely).

Now minimal surfaces are are singled out by the property that their spinor fields are *meromorphic*. This characterization is in fact equivalent to the integrability conditions, which we rewrite as follows: Let $\mathcal{M} = C^\infty(M^2; \widetilde{\mathbf{C}})$ denote the set of functions on $M^2$ taking values in the extended complex plane, and denote the conjugation operator on $\mathcal{M}$ by $\mathcal{C}(z) := \bar{z}$. Then we define $\partial\!\!\!/ := \partial_z \circ \mathcal{C} : \mathcal{M} \to \mathcal{M}$, where $\partial_z = \frac{1}{2}(\partial_x - i\partial_y)$. Observe that $\partial\!\!\!/^2 = \partial_z \partial_{\bar{z}} = \partial_x^2 + \partial_y^2 = \nabla^2$, so this is indeed the Dirac operator on $\mathcal{M}$.

The Dirac equation again reads $(\partial\!\!\!/ + iM)f(z) = 0$, and setting $M = 0$ we obtain the two-dimensional Weyl equation

$$\partial\!\!\!/ \, f = 0 \qquad \Leftrightarrow \qquad \partial_{\bar{z}} f = 0$$

But this is nothing but the Cauchy-Riemann equations![10] Thus the integrability condition for a minimal surface with wavefunction $\psi$ is simply the 2-dimensional version of the equation of motion for a massless spin-1/2 particle

$$\partial\!\!\!/ \, \psi = 0 \qquad \Leftrightarrow \qquad \mathbf{x} \text{ is minimal.}$$

This is the first piece of evidence that the particle/surface analogy may have some content, at least in special cases.

As we have seen, Dirac showed that the behavior of spin is somehow dictated by the equation of motion: the wavefunction's spinoral equivariance is a manifestation of relativistic invariance. Something similar occurs in the case of minimal surfaces: the simple form of the integrability condition $\partial\!\!\!/ \, \psi = 0$ is part of what makes the Goursat transform possible, since the transformation $\psi \to A\psi$ only respects the zero-mass equation. Thus minimal surfaces are the only surfaces that can be *complex* rotated. What does this rotation mean in our physical picture? Complex rotation is not a rotation in space, but in some "extra" dimension, and we argue that if this unseen dimension is interpreted as *time*, a Goursat transformation is nothing but a Lorentz transformation.

To make this interpretation precise, first recall that the Goursat transform of a surface $\mathbf{x} = Re\boldsymbol{\gamma}$ is a complex rotation of the velocity $\dot{\boldsymbol{\gamma}}$ of its associated minimal curve, and this velocity takes values in the null quadric $Q_1$, which is double-covered by the spinor map $\varphi : \mathbf{C}^2 \to Q_1$. The set $\mathcal{L}_+ = \{ X \in Herm_2(\mathbf{C}) \mid \, detX = 0, \, trX \geq 0\}$ is also double covered $\lambda : \mathbf{C}^2 \to \mathcal{L}_+$ by the outer product $\lambda(\psi) = \psi\psi^*$, so that

[10] That the (zero-mass) Dirac equation is a natural 4-dimensional generalization of the Cauchy-Riemann equations is easily one of its most beautifully appealing characteristics.

the composition $\lambda \circ \varphi^{-1}$ is a well-defined bijection. The set $\mathcal{L}_+$ is just the positive light cone in $Herm_2(\mathbf{C})$, which can be identified with the light cone $\mathcal{L}^+ = \{\,\mathbf{x} \in \mathbf{R}^{1,3} \mid \|\mathbf{x}\| = 0,\ x_0 \geq 0\}$ by restricting the Pauli map $S : \mathcal{L}^+ \to \mathcal{L}_+$. Thus the quadric $Q_1$ can be identified with $\mathcal{L}^+$ via the composition

$$
\begin{array}{lllcl}
L \;:=\; S^{-1} \circ \lambda \circ \varphi^{-1} & : & Q_1 & \xrightarrow{\simeq} & \mathcal{L}^+ \\
 & & \begin{pmatrix} \frac{1}{2}(q^2 - p^2) \\ \frac{i}{2}(q^2 + p^2) \\ qp \end{pmatrix} & \longmapsto & \frac{1}{2}(|p|^2 + |\,q|^2,\ 2\,\mathrm{Re}(p\bar{q}),\ 2\,\mathrm{Im}(p\bar{q}),\ |p|^2 - |\,q|^2)
\end{array}
$$

Denoting the 2:1 coverings of $SO_3(\mathbf{C})$ and $SO_{1,3}(\mathbf{R})$ by $\Sigma_1 : SL_2(\mathbf{C}) \to SO_3(\mathbf{C})$ and $\Sigma_2 : SL_2(\mathbf{C}) \to SO_{1,3}(\mathbf{R})$, respectively, the maps

$$
\begin{cases} \varphi : \mathbf{C}^2 \to Q_1 \\ \lambda : \mathbf{C}^2 \to \mathcal{L}_+ \\ S : \mathcal{L}^+ \to \mathcal{L}_+ \end{cases} \qquad \text{satisfy} \qquad \begin{cases} \varphi\,(\Sigma_1(A)\mathbf{x}) = A\,\varphi(\mathbf{x}) \\ \lambda\,(A\mathbf{x}) = A\,\lambda(\mathbf{x})A^* \\ S\,(\Sigma_2(A)\mathbf{x}) = A\,S(\mathbf{x})A^* \end{cases}
$$

and therefore $L\,(\Sigma_1(A)\mathbf{x}) = \Sigma_2(A)\,L(\mathbf{x})$ for all $A \in SL_2(\mathbf{C})$. $L$ induces a group isomorphism

$$
\begin{array}{rcl}
\Lambda \;:\; SO_3(\mathbf{C}) & \to & SO_{1,3}(\mathbf{R}) \\
O & \mapsto & L \circ O \circ L^{-1}
\end{array}
$$

and since $\Sigma_1$ and $\Sigma_2$ are both extensions of $\Sigma : SU_2 \to SO_3(\mathbf{R})$, the equivariance of $L$ implies that $\Lambda$ restricts to the identity map on $SO_3(\mathbf{R})$, in the sense that

$$
\begin{array}{rcl}
\Lambda \;:\; SO_3(\mathbf{R}) & \to & SO_3(\mathbf{R}) \subset SO_{1,3}(\mathbf{R}) \\
R & \mapsto & \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}
\end{array}
$$

Thus a rotation of $\dot{\boldsymbol{\gamma}}$ through a complex angle (a "pure complex rotation," in the sense of Goursat) corresponds to a "boost" of $L(\dot{\boldsymbol{\gamma}})$ in Minkowski space (that is, acting on the time component). This is our physical interpretation the Goursat transformation: When the velocity of the associated minimal curve is regarded as living in Minkowski space, complex rotation *is* Lorentz transformation.[11] That surfaces appear to bend under Lorentz transformation is a reflection of the fact that the time direction has been "projected out" via $Re : \boldsymbol{\gamma} \mapsto \mathbf{x}$ (or, as we have already mentioned, because Lorentz transformation acts by Möbius transformation on the celestial sphere, which does not extend to isometry in Euclidean 3-space).

This constitutes the second piece of evidence that minimal surfaces can be fruitfully thought of as particles, namely, as a *left-handed* spin-1/2 massless relativistic particles. Regarding the Goursat transformation as a Lorentz transformation and the spin curve as a wavefunction, our previous result on the Goursat-equivariance of spin curves of minimal surfaces can be restated to parallel Dirac's result: The Euclidean equivariance of the wavefunction extends to a Lorentzian equivariance, specifically that of a left-handed 2-spinor $\psi(\Sigma(A)\mathbf{x}) = A\,\psi(\mathbf{x})$. Moreover, this equivariance is a consequence of the Dirac equation describing its motion, and in this sense minimal surfaces are the "relativistically invariant" surfaces in $\mathbf{R}^3$.

By now it is apparent, for dimensional reasons, that attempting to regard surfaces in Euclidean 3-space as relativistic is a bit odd; the most natural ambient (flat) space is clearly the usual Minkowski space-time $\mathbf{R}^{1,3}$. The remaining horospherical surfaces in the negatively curved space forms $f : M^2 \to H^3_\epsilon \subset \mathbf{R}^{1,3}$ indeed live in space-time, and as we now argue, constitute a very natural class. Using Bryant's representation theorem, these surfaces also have a concept of spin - a wavefunction $\psi$ such that the structure equations are equivalent to the 2-dimensional Weyl equation, $\bar{\partial\!\!\!/}\,\psi = 0$. We restrict ourselves to the *right-invariant* representation of horospherical surfaces: to say that $f$ has wavefunction $\psi$ means that $f = \pi(F)$, where $F : M^2 \to SL_2(\mathbf{C})$

[11]Conversely, this suggests that doing relativity in the light cone $\mathcal{L}^+$ is equivalent to studying the $SO_{1,3}(\mathbf{R})$ geometry of $Q_1$. Although there is no obvious advantage to doing so, the notion of using $SO_3(\mathbf{C})$ in place of the usual Lorentz group does at least seem philosophically consistent with the twistor programme of Penrose, which aims to make the geometric objects of physics "as holomorphic as possible."

is a holomorphic solution of $dFF^{-1} = \tau$, where $\tau$ is constructed in the usual way from $\psi$. Then as we have seen, the desired left-handed relativistic equivariance of the wavefunction (which requires so much effort to even articulate for minimal surfaces in $\mathbf{R}^3$) is automatic for hyperbolic horospherical surfaces: the wavefunction transforms according to $\psi \to A\,\psi$ under a Lorentz transformation of the surface $f \mapsto \Sigma(A)\,f$. The right-invariant representation also preserves our geometric interpretation of the spin vector, whose direction is now identified with the *hyperbolic* Gauss map $g_h$, since this map is still given by the quotient of the spinor fields $g_h = [\psi] = \frac{p}{q}$.

Recall also that the right-invariant representation is the natural choice when one studies the *dual* correspondence, instead of the Lawson correspondence. This is the rule by which surfaces in different hyperbolic spaces agree are considered *cousins* if their Hopf differentials and hyperbolic Gauss maps agree, and the result is that horospherical surfaces with the same wavefunction $\psi$ are cousins. For the purposes of our analogy, this phenomenon raises a few obvious questions:

1.) What is the physical relationship between surfaces with the same wavefunction? Should they be considered somehow the same particle?

2.) Why should these surfaces be confined to submanifolds ($H^3_\epsilon$) of Minkowski space in the first place? Does this restriction have physical meaning?

A number of reasonable answers seem possible for making sense of these issues; we offer some simple conjecture here. Recall that the Fourier transform $\mathcal{F} : L^2(\mathbf{R}^3) \to L^2(\mathbf{R}^3)$ isometrically interchanges the position and momentum observables $M_{x^i} \leftrightarrow i\hbar\,\partial_i$. Thinking of this as formally renaming $\{t, x^i\}$ with new labels $\{E, p_i\}$, applying the transformation allows us to regard our surfaces as living in momentum space. The advantage of doing so is that the hyperbolic spaces have a natural physical meaning in this setting: they can be thought of as *mass shells* $X_m$ for a particle with mass $m$ (the set of 4-momenta which satisfy the relativistic energy relation)

$$X_m := \left\{ \mathbf{p} \in \mathbf{R}^{1,3} \mid \ \|\mathbf{p}\|^2 = m^2 \right\}$$

It seems natural then to regard a surface in a hyperboloid of radius $m$ as a particle with this mass. Surfaces sharing the same wavefunction might then be regarded as a single accelerating particle.

The final aspect of the analogy that deserves comment is the physical relevance of total curvature. As mentioned above, there is a tendency to regard total curvature as a kind of "charge" of a geometric object in the event that it is quantized. For example, by choosing to model his magnetic monopole on a principle $U(1)$-bundle, Dirac was able to conclude that, if magnetic charge exists, it must be quantized. Mathematically, this is because he identified charge with the total curvature, which was none other than the Chern number, and as a result, Chern numbers of general bundles are sometimes thought of as charges. In string theory, the total Gaussian curvature $\iint K dA$ is already extremely relevant in the compact case since, by the Gauss-Bonnet theorem, it determines the topology. By Osserman's result, the *complete* surfaces have a non-compact analog of Gauss-Bonnet: the total curvature is quantized, and the topology imposes a bound (called the Cohn-Vossen inequality). This fails to hold for the corresponding horospherical surfaces, but certain subclasses of these surfaces instead have quantized *Willmore energy* $\iint (H^2 - K) dA$ (see [3]). In any case, these quantization results strongly invite interpretations like "geometrical charge" for the values of these functionals on horospherical surfaces, and conversely, this phenomenon makes the Willmore energy a potentially relevant physical quantity for string theory, particularly in the non-compact case.

What is the upshot of the particle/surface analogy? Perhaps the best feature of our scheme (and mathematical physics in general) is that it tends to yield interesting conjectures. For example, the well-known process of generalizing the Dirac equation to non-flat space-times (essentially, by replacing partial differentiation with covariant differentiation) may suggest the right method for generalizing the spinor representation to surfaces in an arbitrary ambient Riemannian 3-manifold. Similarly, since a particle's spin quantum number determines the dimension of the vector space (representation) in which its wavefunction takes values, this number may also assist in guessing the codimensions where spinor representations are even possible. Our reasoning also suggests a general programme for surface theory by which representation theorems like those of Weierstrass and Bryant are sought out on the basis of their analogy to the abundant laws and equations found in particle physics. We again emphasize that, like much of the work currently considered mathematical physics, all of our observations here have been based on purely formal analogy, not in any way on physical evidence or observation, and such analogies are usually only beneficial to mathematicians, providing new ways to regard the abstract objects of their study.[12] In the author's estimation, such remarks will likely have little consequence in actual physical theory, except as general endorsement of differential geometry as a good context for its foundation.

[12] Dirac's achievements, indeed many of the greatest accomplishments of mathematics and physics, are a testament to the somewhat reckless imagination involved in this sort of formal reasoning. There appears to be a wealth of hidden beauty in the universe attainable by those willing to seriously consider the implications of ideas which at first seem preposterous.

# Bibliography


[1] W. Blaschke. *Affine Differentialgeometrie*. Springer-Verlag, Berlin, 1923.

[2] D. Bleecker. *Gauge Theory and Variational Principles*. Dover Publications, 1981.

[3] C. Bohle and G.P. Peters. Bryant surfaces with smooth ends. *arXiv:math.DG/0411480*, 2005.

[4] R. Bryant. Surfaces of mean curvature one in hyperbolic space. *Asterisque*, 154-155, 1987.

[5] F.E. Burstall, D. Ferus, K. Leschke, F. Pedit, and U. Pinkall. *Conformal Geometry of Surfaces in $S^4$ and Quaternions*. Springer, 2001.

[6] E. Cartan. *La theorie des groupes finis et continus et la geometrie differentielle traitees par la methode du rep‘ere mobile*. Gauthier-Villars, Paris, 1937.

[7] P.A.M. Dirac. *The Principles of Quantum Mechanics*. Clarendon Press, Oxford, 1930.

[8] C.F. Gauss. *General investigations of curved surfaces*. Raven Press, Hewlett, NY, 1965.

[9] P.R. Girard. The quanternions and modern physics. *Comm. Dublin Inst. Adv. Studies*, 1984.

[10] E. Goursat. Sur un mode de transformation des surfaces minima. *Acta Math.*, 11, 1887-8.

[11] U. Hertrich-Jeromin. *Introduction to Mobius Differential Geometry.* London Math. Soc. Lecture Note Series 300. Cambridge U. Press, 2003.

[12] G. Jensen, E. Musso, and L. Nicolodi. *Surfaces by Moving Frames.* in preparation, 2009.

[13] J-L. Lagrange. *Oeuvres de Lagrange.* Gauthier-Villars, Paris, 1867-92.

[14] B. Lawson and M-L. Michelsohn. *Spin Geometry.* Princeton University Press, 1989.

[15] J-B. Meusnier. Memoire sur la courbure des surfaces. *Memoires des Savants etrangers*, X, 1785.

[16] R. Osserman. *A survey of minimal surfaces.* Cambridge Univ. Press, New York, 1989.

[17] J. Perez and A. Ros. Properly embedded minimal surfaces with finite total curvature. In G. Pirola, editor, *The global theory of minimal surfaces in flat spaces, Lecture Notes in Math., 1775.* Springer, 2002.

[18] J.G. Ratcliffe. *Foundations of Hyperbolic Manifolds.* Springer-Verlag, New York, 2006.

[19] J.L. Synge. Quanternions, lorentz transformations, and the conway-dirac-eddington matrices. *Comm. Dublin Inst. Adv. Studies*, 1972.

[20] M. Umehara and K. Yamada. Complete surfaces of constant mean curvature 1 in the hyperbolic 3-space. *Annals of Math.*, 137, 1993.

[21] M. Umehara and K. Yamada. A duality on cmc1 surfaces in hyperbolic space, and a hyperbolic analogue of the osserman inequality. *Tsukuba J. of Math.*, 21, 1997.

[22] K. Voss. Uber vollständige minimalflächen. *L'Enseignement Math.*, 1964.

[23] K. Weierstrass. Uber eine besondere gattung von minimalflächen. *Monatsberichte der Berliner Akademie*, 1867.