

# A Proof That Witten's Open String Theory Gives a Single Cover of Moduli Space

**Barton Zwiebach**\*

Center for Theoretical Physics, Laboratory for Nuclear Science and Department of Physics,  
Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received December 30, 1990; in revised form March 29, 1991

**Abstract.** We show that Witten's open string diagrams are surfaces with metrics of minimal area under the condition that all nontrivial open Jordan curves be longer or equal to  $\pi$ . The minimal area property is used together with a mini-max problem to establish a new existence and uniqueness theorem for quadratic differentials in open Riemann surfaces with or without punctures on the boundaries. This theorem implies that the Feynman rules of open string theory give a single cover of the moduli of open Riemann surfaces.

## 1. Introduction and Summary

The open string field theory proposed by Witten [Wi] is based on a single interaction joining three open strings. This interaction gives a simple set of Feynman rules that build string diagrams. A string diagram is a Riemann surface with a choice of analytic coordinates at punctures. In open string theory the Riemann surfaces are surfaces with boundaries, and the punctures, that must lie on the boundary components, represent the external open strings. Typically, if we endow a Riemann surface with a suitable metric, it is possible to extract analytic coordinates around the punctures. This is the case for the string diagrams of [Wi], which are conveniently described in terms of metrics.

A necessary condition for the consistency of open string theory is that the Feynman rules construct string diagrams that provide a single cover of all relevant moduli spaces. While there is a fair amount of evidence [Wi, GMW, Gi, Og, BS, Sa], and it is widely believed to be the case, we lack a complete proof. This important point has been correctly emphasized by Samuel [Sa]. Not only we are missing a complete proof, but also a proper understanding of why it works. As we have learned in closed string theory, the fact that a set of Feynman rules work well at the classical level, namely, produce a single cover of moduli spaces of

---

\* Supported in part by funds provided by the U.S. Department of Energy (D.O.E.) under contract # DE-AC02-76ER03069

genus zero punctured surfaces, does not guarantee that they will work for higher genus surfaces. Thus gauge invariance of the classical theory does not explain why Witten's theory gives a single cover of moduli space to all genus. Physically one expects correct covering to all genus if the path integral measure is BRST invariant. It is not yet clear, however, whether or not the path integral measure of [Wi] is BRST invariant [Th].

One way to understand why string diagrams give a correct covering of moduli space is to find a principle that defines for any given Riemann surface a unique string diagram. Such principle guarantees that the set of all inequivalent string diagrams give a single cover of moduli space. Overcounting cannot happen because it would imply that there exists more than one string diagram for a given Riemann surface, in contradiction with uniqueness. No surface can be missed since for any surface there is a string diagram, which must therefore be included in the set of all string diagrams. It must also be checked that the Feynman rules construct all string diagrams precisely once. All in all, this shows that the Feynman rules give a single cover of moduli space.

In this sense, light-cone diagrams work because they are the unique solution to the problem of finding a meromorphic abelian differential of specified residues and with pure imaginary periods [GW]. Similarly, the polyhedral diagrams [SaZw, KKS] of classical closed string field theory work because they are the unique solution of a problem that asks for the metric of minimal area under the condition that any nontrivial closed curve be longer or equal to  $2\pi$  [Zw1]. This same minimal area problem seems to define the string diagrams and Feynman rules of the full quantum closed string theory [Zw2].

In this paper we show that the open string diagrams of Witten are also the unique solutions of a new minimal area problem. The minimal area problem asks for the *metric of minimal area under the condition that all nontrivial Jordan open curves (curves without self intersections) be longer than  $\pi$* . In fact it will be quite straightforward to show this is the case. This property explains why Witten's string diagrams do not overcount. In order to show that no surface is missed we have to show existence, namely, that this minimal area problem always has a solution, which, in fact, is a Witten digram. This second part is much harder to prove but turns out to be very instructive. It is almost a miracle that *the solution of this minimal area problem is always given by a quadratic differential*. It is well known that quadratic differentials arise from minimal area problems in which we impose length conditions on curves homotopic to a set of nonintersecting Jordan curves [St]. Nevertheless the set of all nontrivial Jordan open curves in a surface is an infinite set of curves with intersections! The proof explains why the metric arises from a quadratic differential, and illuminates why dealing with open curves is much simpler than dealing with closed ones. The end result is a new existence and uniqueness theorem for quadratic differentials. This theorem implies that the Feynman rule of open string field theory give a single cover of moduli space.

The earlier work on the open string diagrams of [Wi] began with Ref. [GMW] who elucidated to a large degree the structure of the string diagrams, and gave an argument to show that one has a single cover for Riemann surfaces one boundary and no punctures. This argument was subsequently extended in [Gi] to the case of several boundary components and no punctures. Actually these particular results follow directly from Strebel's existence and uniqueness theorem for quadratic differentials of specified heights [St]. We will explain how this happens, and in

this way simplify the discussion of [GMW, Gi]. For the case of surfaces with punctures at the boundary, the nature of the problem is somewhat different and no known theorem seems to apply.<sup>1</sup> This is actually not surprising. Quadratic differentials on open surfaces are typically studied by doubling. Open string tree level diagrams when doubled give closed string diagrams, whose quadratic differentials have closed trajectories and both punctured disks and intermediate annuli. The existence and uniqueness theorem for quadratic differentials with both punctured disks and annuli, was only established recently [Zw1] by combining earlier theorems of Jenkins and Strebel. While there are no complete proofs, two kind of arguments have been given that suggest strongly that things ought to work for surfaces with boundaries and punctures. In the first on [Gi], one argues that the vacuum diagrams define string diagrams with punctures by factorization; it seems possible to turn this into a rigorous argument for the existence of an open string diagram for any punctured surface (although uniqueness would not follow immediately). In the second one [BS], one argues that BRST invariance is clear since all Feynman graphs match properly and produce correct singularities in the open string sector.

We will assume some familiarity with quadratic differentials and with [Zw1]. Background material for the present paper can be found in [Zw3], where we give some simple theorems relating homotopy classes of open and closed curves, and discuss the issue of regularization of the area for open surfaces with punctures. The necessary results will be quoted. In that work we discuss a different minimal area problem for open string theory, one yielding at the quantum level string diagrams that are manifestly factorizable both in open and closed string channels, and that lead to an action that satisfies manifestly the full Batalin-Vilkovisky master equation [Zw4] (the open string diagrams of [Wi] do not satisfy those properties).

The problem studied here defining open string diagrams can also define string diagrams when there are closed strings present in the open Riemann surface. This leads to a remarkable simple way of introducing closed strings into the open string formalism of Witten, giving a single copy of moduli space. In contrast with the theory proposed in [Zw4], this latter theory does not seem to be as fundamental since the closed strings must be on-shell. As a way to generate moduli space, however, the Feynman rules are as simple as can be. These results will be presented in [Zw5].

*Brief Sketch of the Results.* Our basic result is an existence and uniqueness theorem for a special type of quadratic differentials on Riemann surfaces with at least one boundary component. These quadratic differentials, denoted as open string quadratic differentials are defined as follows.

**Definition.** An *open string quadratic differential* on a Riemann surface  $R$  with at least one boundary component, and possibly punctures on the boundaries is a quadratic differential whose singularities are zeroes, and second order poles at the punctures. Its characteristic domains are rectangular strips, finite or semi-infinite. The strips are covered by open horizontal trajectories of length  $\pi$ . The boundaries of the strips are horizontal edges, corresponding to critical horizontal trajectories (two or one if the strips are finite or semi-infinite, respectively) and two vertical

<sup>1</sup> As a consequence, the commonly made claim that open string theory gave a cell decomposition that was previously known appears to be incorrect

trajectories. The strips are joined together along their horizontal edges with all zeroes of the quadratic differential at the midpoint of these edges. The endpoints of a horizontal edge are never identified. The vertical edges of the strips are left open and make up the boundary of  $R$ . (The double of an open string quadratic differential is a JS quadratic differential).

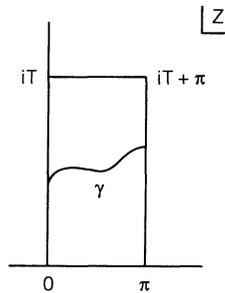
Our main result is a theorem establishing the existence and uniqueness of open string quadratic differentials and their relation to a minimal area problem:

**Theorem 4.** *Given a Riemann surface  $R$  with at least one boundary component and possibly punctures on the boundary (except for the disk with one or no punctures) there is a unique open string quadratic differential on  $R$ . This quadratic differential gives rise to the metric of minimal area under the condition that all nontrivial Jordan open curves on  $R$  be longer or equal to  $\pi$ .*

A simple consequence of this theorem is the desired physical result establishing that.

**Theorem 5.** *The Feynman diagrams of open string theory give a single cover of the moduli spaces of Riemann surfaces with  $b \geq 1$  boundaries and  $m \geq 0$  punctures.*

*Brief Sketch of the Proof.* Most of our work in the present paper is devoted to make the following arguments precise. It is simple to understand why the open string diagrams solve the advertised minimal area problem. Given a flat strip of height  $T$  and width  $\pi$ , as shown in Fig. 1, the flat metric  $g = 1$  (recall  $dl = g|dz|$ ) is actually the metric of minimal area under the condition that any curve  $\gamma$  with



**Fig. 1.** A rectangular region in the  $z$ -plane, of parameter height  $T$  and parameter width  $\pi$ . The flat metric  $g = 1$  gives the minimal area metric under the condition that any curve  $\gamma$  with endpoints on the vertical segments be longer or equal to  $\pi$

endpoints at the two vertical edges be longer or equal to  $\pi$ . But in fact the open string diagram is made up of strips. It can be shown that in open string diagrams, on every strip, open curves going from one boundary to the other, are nontrivial. Under the length condition on those curves alone, the string diagram gives the least possible area. Thus one only has to verify that all other nontrivial open curves are longer than  $\pi$ , and this is fairly obvious.

In order to tackle the existence of the string diagram, we first study a class of quadratic differentials that are mirror symmetric. Throughout this paper  $R$  will denote an open Riemann surface,  $R^*$  will be its mirror image, and  $\hat{R}$  will be the doubled surface, made by gluing  $R$  and  $R^*$  across their boundaries. We choose an

admissible set of open curves in  $R$  (nontrivial Jordan curves that do not intersect nor are homotopic to each other). By doubling we get a mirror symmetric admissible set of closed curves in  $\hat{R}$ . We pose the problem of minimizing the area of  $\hat{R}$  under the condition that any curve homotopic to one in the admissible set be longer or equal to  $2\pi$ . The answer is a metric arising from a Jenkins-Strebel quadratic differential having both punctured disks and internal annuli [Zw1]. It is nice that the closed curves in  $\hat{R}$  representing the boundary components of  $R$  appear as geodesics on the metric defined by the quadratic differential. Due to the mirror symmetry, in each cylinder the boundary components of  $R$  appear as vertical trajectories of the quadratic differential splitting each cylinder into two rectangular strips of width  $\pi$ . This shows that  $R$  is built by joining strips. But this is not enough in order to have an open string diagram, and at any rate this could not be the answer since there is an infinite number of choices for the open curves on  $R$ , yielding many (or possibly infinite) different quadratic differentials.

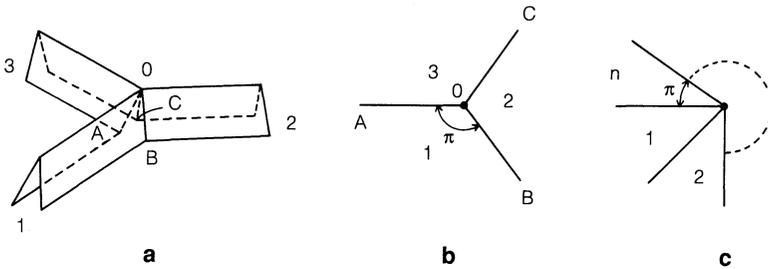
The solution is to try all possible choice of curves in  $R$ , get all possible quadratic differentials, and choose the one of largest area. It will be shown that this quadratic differential is unique, and yields an open string diagram. Choosing the quadratic differential of largest area guarantees that any nontrivial open curve homotopic to a critical trajectory must be longer than  $\pi$ . Otherwise, the double of the curve could be added to the admissible set, and we would find another quadratic differential of even larger area. This length condition on open curves is very powerful. It allows us to rule out the possibility that the boundary of  $R$  lies on a critical (horizontal) trajectory. This is simply because on each boundary of a ring domain we already have two points corresponding to boundaries of  $R$ , any other piece of  $R$  appearing on this critical trajectory will give rise to a short nontrivial open critical curve joining the two boundaries, which is not permitted. It also allows us to show that the boundaries of  $R$  cannot hit zeroes of the quadratic differential. The patterns of gluing of the horizontal edges of the open strips are therefore severely limited. For example there cannot be self-identification of segments, because we would get a critical open curve shorter than  $\pi$ . A little thought shows that only the symmetric gluing pattern of open string diagrams, with zeroes at the midpoint of the horizontal segments is allowed.

## 2. Open String Diagrams as Minimal Area Surfaces

In this section we begin by describing the Feynman diagrams of the open string field theory [W]. We think of a Feynman diagram as a collection of data (lengths of propagators, and parameters of the graph) which determine a Riemann surface. Each Feynman diagram defines a surface with a canonical metric, and a canonical quadratic differential, which will be denoted as an *open string quadratic differential*. These quadratic differentials were discussed in [GMW]. Moreover given a surface with an open string quadratic differential we can canonically obtain a Feynman diagram. The above define maps between the set of inequivalent Feynman graphs and the set of inequivalent open string quadratic differentials. We will see that the maps define an isomorphism between these two sets. As explained in detail in Sect. 2.1 of [Zw3], this fact, together with an existence and uniqueness proof for open string quadratic differentials, to be proven later, guarantees the single cover of moduli space. We also show that the open string diagrams are surfaces with a metric of minimal area under

the condition that all nontrivial Jordan open curves be longer or equal to  $\pi$ . A nontrivial open curve is one that cannot be continuously shrunk away while keeping its endpoints on the boundary components of the surface. The endpoints cannot be moved across punctures lying on the boundary (for more details on nontrivial open curves see [Zw3]). The minimal area property is actually simple to prove, and it is useful because, even without the existence and uniqueness theorem, it already guarantees that there cannot be overcounting. Moreover, we will also use it to prove the uniqueness part of the existence and uniqueness theorem for open string quadratic differentials.

Let us begin with the Feynman diagrams. The covariant open string field theory of [Wi] is based on a single three string vertex and a propagator that consists of flat rectangular strips of width  $\pi$  and length  $T$  with  $T \in [0, \infty]$ . The vertex joins three such strips symmetrically creating a singularity of total internal angle equal to  $3\pi$  at the midpoint, as illustrated in Fig. 2a, b. While this is the only

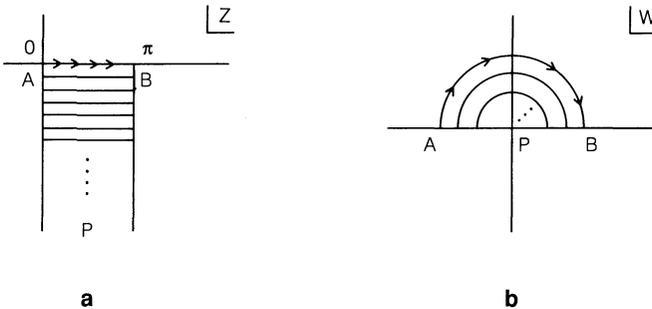


**Fig. 2. a** The Witten vertex for three open strings joins symmetrically three strips of width  $\pi$ . The strips, labeled 1, 2 and 3 are glued via the segments  $AOB$ ,  $BOC$  and  $COA$  which correspond to critical horizontal trajectories. Point  $O$  is the midpoint of these segments. **b** The total internal angle at  $O$  is  $3\pi$  since every strip contributes an angle  $\pi$ . Point  $O$  is a first order zero of a quadratic differential. **c** When  $n$  strips are joined symmetrically we have a zero of the quadratic differential of order  $(n - 2)$

elementary vertex, when propagators joining different vertices collapse ( $T = 0$ ), we effectively get situations where more than three strings join symmetrically. Thus in all generality, the string diagrams include configurations in which  $n \geq 3$  strings join as illustrated in Fig. 2c. Here the singularity created at the midpoint corresponds to a neighborhood of total internal angle  $n\pi$ . External open string states are represented by semi-infinite strips. We will always think of the strips as a rectangular region in the  $z$  plane with corners  $z = 0, \pi, iT$  and  $iT + \pi$  (Fig. 1). The horizontal edges, from  $z = 0$  to  $z = \pi$  and from  $z = iT$  to  $z = iT + \pi$  are to be glued to other strips using the gluing patterns of Fig. 2. The vertical edges are left open and correspond to the boundary components of the surface. The Feynman diagram is the pair  $(R, D)$ , where  $D$  is the data concerning the propagator times, and the way they are joined to each other, and  $R$  is the resulting Riemann surface. To each Feynman diagram we can find a string diagram  $(R, \varrho)$  (a surface with a metric, expected to give local coordinates) by defining  $\varrho = 1$  (recall  $dl = \varrho|dz|$ ) on every strip. This is just the flat metric.

In fact this metric arises from a quadratic differential. We can therefore associate canonically to each Feynman diagram a string diagram  $(R, \varphi)$ , as follows: on every finite strip, such as that of Fig. 1, the quadratic differential is defined of the form  $\varphi = \phi(z) (dz)^2 = (dz)^2$ . Thus, the vertical edges correspond to vertical

trajectories of the quadratic differential ( $\varphi$  is real and negative). The horizontal edges of the strip are horizontal trajectories ( $\varphi$  is real and positive). It is possible to glue the strips as indicated in Fig.2 and get a quadratic differential on the whole surface. The singularities are the zeroes of the quadratic differential, and the order of a zero is given by  $(n - 2)$ , where  $n$  is the number of strips that are glued together. Indeed, the familiar pattern of three strips joining symmetrically corresponds to a first order zero of a quadratic differential, with the correct value of  $3\pi$  for the total angle. Therefore the horizontal edges of the strips are critical horizontal trajectories (horizontal trajectories that go through critical points). For the semi-infinite strips, all of the above holds, the only difference is that there is just one horizontal edge. The horizontal edge can be chosen to be the horizontal line joining  $z = 0$  to  $z = \pi$ , and the vertical edges go from this point down the imaginary axis all the way down, as shown in Fig. 3a.



**Fig. 3. a** A semi-infinite strip of width  $\pi$ . The horizontal line extending from zero to  $\pi$  (shown with arrows) is a critical horizontal trajectory. **b** This strip can be mapped to a half-disk, the point at  $w = 0$  corresponds to the puncture

When the string diagram is built, horizontal edges become critical trajectories, and zeroes appear at  $z = \pi/2$ . Note that the strip can be mapped in a unique way into the half disk  $|w| \leq 1, \text{Im}(w) \geq 0$  (Fig. 3b) via  $w = -\exp(-iz)$ . The  $w$  coordinates are the standard analytic coordinates around the puncture ( $w = 0$ ). This is the way to extract coordinates from the flat metric, or the quadratic differential on the surface. The puncture corresponds to a second order pole of the quadratic differential, since one verifies that in  $w$  coordinates (where the puncture is at a finite point) the quadratic differential is  $\varphi = -(dw)^2/w^2$ . The horizontal trajectories in the above quadratic differential are all of length  $\pi$  and have boundary endpoints. If we double the surface, and extend the quadratic differential to the double (this can be done since the quadratic differential is real along the boundaries), we obtain a quadratic differential with closed horizontal trajectories of length  $2\pi$ . Each finite strip becomes a finite cylinder, or annulus, and every semi-infinite strip becomes a semi-infinite cylinder, or a punctured disk, with a second order pole at the puncture. All of this establishes that we have an open string quadratic differential (see the definition in Sect. 1).

Our previous discussion shown that there is a map  $\psi$  from the set  $\mathcal{F}$  of inequivalent Feynman diagrams of open string theory into the set  $\mathcal{S}$  of inequivalent string diagrams  $(R, \varphi)$ , where  $\varphi$  is an open string quadratic differential. It is also clear that an open string quadratic differential defines a unique Feynman

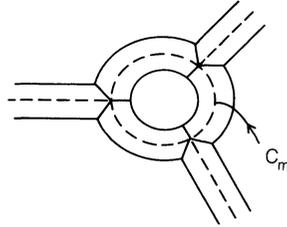
diagram; the horizontal foliation determines for us all the strips. To each finite strip we associate a propagator of length equal to the height of the strip, and to each semi-infinite strip we associate an external state. This defines a map  $\psi^{-1} : \mathcal{S} \mapsto \mathcal{F}$ . By their explicit definition it follows that  $\psi$  and  $\psi^{-1}$  are inverses of each other. We therefore have the desired isomorphism between the sets  $\mathcal{F}$  and  $\mathcal{S}$ . *The Feynman diagrams produce a single copy of all possible open string quadratic differentials.* Two different Feynman diagrams produce two different open string diagrams.

Having established the desired isomorphism, let us turn to the minimal area property. We first need some results on the open string diagrams. There is a useful curve  $\mathcal{C}_m$  in open string diagrams [GMW]. On every strip it runs parallel to the boundary in the middle of the strip cutting it into two equal halves. It has self-intersections at every vertex, and goes to infinity in every semi-infinite strip.

**Lemma 1.** *On any open string diagram the following properties hold:*

- (i) *If the string diagram is cut along the curve  $\mathcal{C}_m$  we obtain a cylinder of height  $\pi/2$  for each boundary component without punctures, and  $m$  infinite strips of height  $\pi/2$  for each boundary component with  $m$  punctures.*
- (ii) *On every strip in the string diagram a Jordan open curve joining the two boundaries of the strip is nontrivial.*
- (iii) *Any nontrivial open curve in the open string diagram is longer or equal to  $\pi$ .*

*Proof.* (i) The result is clear and we refer the reader to Fig. 4 to understand the meaning of the statement. (ii) The open curve is homotopic to an open horizontal trajectories  $\gamma$  on the strip. The double of  $\gamma$  on the doubled string diagram is a core curve in a ring domain. Such core curve in a meromorphic JS quadratic



**Fig. 4.** A string diagram representing a surface with two boundary components, one with no punctures and the other with three punctures. We also show the curve  $\mathcal{C}_m$  cutting each strip in the middle. The cut surface includes one cylinder of height  $\pi/2$  and three infinite strips of height  $\pi/2$

differential is always nontrivial (it represents one of the curves in the admissible set of curves, see [St]). If the double of  $\gamma$  is nontrivial,  $\gamma$  has to be a nontrivial open curve (Lemma 6, in [Zw3]). Therefore the original open curve is nontrivial. (iii) Any open curve that does not cross  $\mathcal{C}_m$  must be trivial. This is clear because of (i). If the open curve does not cross  $\mathcal{C}_m$  it begins and ends on the same boundary component. It must lie completely in one of the cylinders or infinite strips of (i). It can therefore be shrunk away. As a consequence any nontrivial open curve must cross  $\mathcal{C}_m$ . Since the shortest distance between a boundary point and  $\mathcal{C}_m$  is  $\pi/2$ , this shows all nontrivial curves must be longer or equal to  $\pi$ .  $\square$

The above properties will allow us to show that the open string diagrams do solve the following minimal area problem:

**Theorem 1.** *An open string diagram for a Riemann surface with boundaries and  $m > 0$  punctures on the boundaries gives the metric of minimal area under the condition that all nontrivial Jordan open curves are longer or equal to  $\pi$ .*

*Proof.* Because of Lemma 1, (iii), we know that on the open string diagram all nontrivial Jordan open curves are longer or equal to  $\pi$ . The length conditions are therefore satisfied and we only have to show that the area cannot be made smaller. Imagine the surface cut into the strips of the open string diagram. If there is a metric of lower area it must have lower area at least on one of the strips. Nevertheless this is impossible. On any strip the curves going from one boundary to the other are nontrivial (Lemma 1, (ii)) and must be longer than  $\pi$ . But under that condition the original flat metric on the strip is already of minimal area, so its area cannot be lowered.  $\square$

Having explained why the open string diagrams are minimal area metrics let us show why this already guarantees that no overcounting takes place. We first need the following lemma.

**Lemma 2.** *If two open string quadratic differentials on a surface  $R$  give rise to the same metric they must be identical.*

*Proof.* Suppose two quadratic differentials  $\phi_1$  and  $\phi_2$  give rise to the same metric:  $|\phi_1(z)| = |\phi_2(z)|$ . Then the two quadratic differentials can only differ by a phase  $\phi_1(z) = \exp(i\theta(z))\phi_2(z)$ . But then this phase can only be a constant. The constant phase can be evaluated along a boundary component of  $R$ . There both quadratic differentials must have the same phase, since both must have a vertical trajectory. Thus the two quadratic differentials must be identical. This establishes the lemma.

It follows that two different Feynman diagrams, which construct different open string quadratic differentials, must construct different metrics. By uniqueness of minimal area metrics [St,Zw1], two different metrics must correspond to two different surfaces. As a consequence, two different Feynman diagrams cannot yield the same surface, and there is no overcounting.

### 3. Vacuum Graphs From QDS of Specified Heights

The discussion in Sect. 2 guarantees that the open string diagrams do not overcount but we still have not shown that they do not miss some surfaces. In order to show this we have to establish that for any surface there is such a string diagram. For surfaces without punctures this is guaranteed by Strebel's height problem. We look at this case next. The case of surfaces without puncture was analyzed explicitly in [GMW], and the arguments presented there are reminiscent of the proof given in [St].

It was shown by Strebel [St, p.107, Theorem 21.1] that given a Riemann surface with an admissible set of curves on it (a set of nontrivial, nonintersecting and nonhomotopic simple Jordan closed curves) the surface can be built in a unique way by gluing together cylinders of specified heights, one for each curve in the admissible set. The circumferences of the cylinders and the lengths of gluing segments give the modular parameters. This theorem applies both the surfaces with and without boundaries [St, p. 116]. The case of a disk is excluded.

For the case under consideration, namely, of a surface with boundaries and no punctures we pick an admissible set of curves as follows. For each boundary

component we pick a closed curve homotopic to it. We therefore get as many curves as boundary components of the surface, except for the case when the surface is an annulus, where we get just one curve. These curves are nontrivial, can be chosen to be nonintersecting, and are nonhomotopic, therefore they make an admissible set. The height condition is simple, we demand that all ring domains have heights equal to  $\pi/2$ . In each cylinder, one boundary represents a boundary component of the Riemann surface and the other boundary is glued with similar boundaries in the other cylinders. Whenever we glue two such boundaries across a gluing segment, we get a strip of total width  $\pi$  and height equal to the length of the gluing segment. Note that the singularities induced by the gluing happen on the midpoint of the strips, namely, a distance  $\pi/2$  away from the boundaries of the surface. The end result is a surface built with strips of width  $\pi$  in the way described in Sect. 2. This shows the existence of open string diagrams for Riemann surfaces with boundaries but no punctures.

#### 4. A Class of Mirror Symmetric QDS

Let us now begin our analysis of existence of open string quadratic differentials, as we defined them in Sect. 1. The first step will be to study an interesting class of quadratic differentials on Riemann surfaces with boundaries. We will obtain those from mirror symmetric admissible systems of curves on the doubled surface. It will be shown that these quadratic differentials are mirror symmetric, and satisfy a number of properties. In Sect. 5 we will show how to choose a special quadratic differential from this class. That special quadratic differential will be shown to be an open string quadratic differential in Sect. 6.

Consider a Riemann surface  $R$  with boundaries and possibly punctures on the boundaries, its mirror image  $R^*$  and its double  $\hat{R}$ . A *mirror symmetric admissible set* in  $\hat{R}$  includes a closed curve surrounding each puncture and a set of closed curves  $\hat{\gamma}_i$ . These closed curves are obtained by doubling a set of open curves  $\gamma_i$  in  $R$ . The set of curves  $\gamma_i$  is a set of nonintersecting, nontrivial homotopy simple Jordan open curves, not homotopic to punctures, nor to each other. Recall that an open Jordan curve has no self-intersections and only its endpoints lie on boundary components.

Let us verify that the admissible sets of curves in  $\hat{R}$  are consistently defined. This requires that all the closed curves be nonintersecting simple Jordan closed curves of nontrivial homotopy not homotopic to each other. All the closed curves are clearly simple Jordan closed curves, and do not intersect each other. They are of nontrivial homotopy since they are made by doubling nontrivial open curves, and cannot be homotopic to each other (see [Zw3, Lemmas 6 and 8]).

For each mirror symmetric admissible set we can now pose a minimal area problem. We demand least (reduced) area under the condition that any curve homotopic to one in the admissible set be longer or equal to  $2\pi$ . As shown in [Zw1], the minimal area metric arises from a Jenkins-Strebel quadratic differential whose characteristic ring domains include both annuli and punctured disks. Every closed horizontal trajectory is of length  $2\pi$  and the quadratic differential has second order poles at the punctures. We shall define this to be the quadratic differential based on a mirror symmetric admissible set of curves. Such quadratic differentials are studied next.

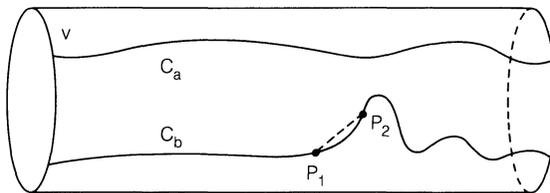
**Theorem 2.** *The following properties hold for a quadratic differential  $\phi$  based on the mirror symmetric admissible set of curves in  $\hat{R}$ :*

- (i) It induces on  $\hat{R}$  a metric invariant under the anticonformal map exchanging  $R$  and  $R^*$  in  $\hat{R}$ .
- (ii) It induces on the open surface  $R$  a quadratic differential for which the boundary components of  $R$  are real trajectories. On every ring domain in  $\hat{R}$  the boundary components of  $R$  appear as vertical trajectories cutting the ring into two identical rectangular strips of height  $\pi$ .
- (iii) The boundary of  $R$  as it appears in a ring domain cannot hit the boundary of the ring domain and become immediately part of that boundary.

*Proof.* (i) If the metric  $\varrho$  were not invariant, the anticonformal map would yield a new metric  $\varrho^*$  on  $\hat{R}$  of the same total area. We claim that the metric  $\varrho^*$  satisfies the same length conditions as the metric  $\varrho$  did. Consider any closed curve  $\gamma$  homotopic to the closed curve  $\hat{\gamma}_i$  of the admissible set:  $\gamma \approx \hat{\gamma}_i$ . The length  $l_\gamma(\varrho^*)$  of  $\gamma$  in the metric  $\varrho^*$ , is clearly equal to the length  $l_{\gamma^*}(\varrho)$ , of  $\gamma^*$  (the image of  $\gamma$ ) in the metric  $\varrho$ . Since  $\gamma \approx \hat{\gamma}_i$ , it follows that  $\gamma^* \approx \hat{\gamma}_i^*$ , however, the latter curve, by construction is invariant under the map, and as a consequence  $\gamma^* \approx \hat{\gamma}_i$ . It therefore follows that  $l_\gamma(\varrho^*) = l_{\gamma^*}(\varrho) \geq 2\pi$ , and indeed the metric  $\varrho^*$  is seen to satisfy the length conditions of the minimal area problem. Since the area of  $\hat{R}$  is the same, and minimal, with both metrics, the uniqueness of minimal area metrics implies  $\varrho = \varrho^*$ .  $\square$

(ii) Any quadratic differential in  $\Phi_{\hat{R}}$  will have semi-infinite tubes of circumference  $2\pi$  at the punctures, and internal annuli corresponding to finite cylinders of circumference  $2\pi$  whose core curves are homotopic to the curves in the specified admissible set of curves (as usual, all annuli may not be realized, since some may collapse). Our first step will be understanding how the boundary components of  $R$ , which are a disjoint set of closed curves  $\mathcal{C}_i$  in  $\hat{R}$ , appear on the cylinders of the quadratic differential.

Any curve in the admissible set can clearly be chosen so that it intersects the set of curves  $\mathcal{C}_i$  at only two points. For a curve homotopic to a puncture, the two intersection points lie on the same component of  $\mathcal{C}_i$ . For some other closed curves the intersection points may lie in different components of  $\mathcal{C}_i$ . It follows that any curve homotopic to one in an admissible set must intersect the relevant boundary component(s) at least two times. Let us consider one of the cylinders of the quadratic differential. Since any closed curve in it must intersect the curves  $\mathcal{C}_i$  in two points at least, the relevant pieces from  $\mathcal{C}_i$  must appear as two curves, each one extending from one boundary of the cylinder to the other, as shown in Fig. 5. The two curves, denoted as  $\mathcal{C}_a$  and  $\mathcal{C}_b$  in the figure, cannot intersect each other and do not have self intersections (since they do not that on  $\hat{R}$ ).



**Fig. 5.** A ring domain in a quadratic differential arising from a mirror symmetric admissible set of curves on the double  $\hat{R}$  of  $R$ . The curves  $\mathcal{C}_a$  and  $\mathcal{C}_b$  are boundary components of  $R$ , they will be shown to be geodesics

Let us first show that the curves  $\mathcal{C}_a$  and  $\mathcal{C}_b$  must be geodesics of the flat metric in the cylinder. If these curves are not locally shortest then there exist two points  $P_1$  and  $P_2$  on the curve through which there is a straight segment shorter than the length of the curve in between these two points (Fig. 5). This shortest segment is unique. Its image under the anticonformal map relating  $R$  and  $R^*$  in  $\hat{R}$  must yield another segment of the same length between the same two points. The uniqueness of the segment implies that the segment must coincide with the boundary component of  $R$ . Thus the boundary of  $R$  is locally shortest, namely, a geodesic. It is also necessary to show that no additional segment  $\mathcal{C}_c$  of a boundary component can appear on this cylinder. If it did, it must also be a geodesic and cannot intersect  $\mathcal{C}_a$  or  $\mathcal{C}_b$ , so it must also go from one boundary to the other one on the cylinder. In fact such extra boundary segments, if they would appear at all, would have to appear in pairs (so that they would cut the cylinder into connected pieces that belong alternatively in  $R$  and  $R^*$ ). Consider Fig. 6, where we show the three boundary segments, and an open curve, which is

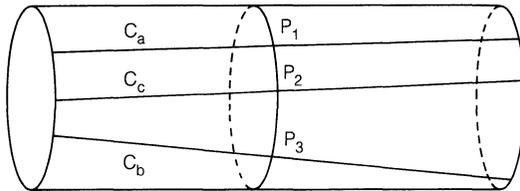


Fig. 6. We need to rule out the possibility that an additional curve  $\mathcal{C}_c$  representing a boundary component of  $R$  could appear in the cylinder

part of a core geodesic on the cylinder, going through  $P_1, P_2$  and  $P_3$ . Under the antiholomorphic map we should get another open curve of the same length and going through the three fixed points  $P_i$ . But this is impossible since near  $P_1$  and  $P_3$  the map must look like a reflection, and the new curve cannot go through  $P_2$ . This is a contradiction, and shows that *only two boundary segments can appear on a cylinder*.

Let us now see why any geodesic core curve in the cylinder must be divided into two pieces, each of length  $\pi$ , by the boundary segments. The situation is illustrated in Fig. 7, where we see the two boundary segments  $\mathcal{C}_a$  and  $\mathcal{C}_b$  dividing the cylinder into two regions, one in  $R$  and the other in  $R^*$ . Consider now a core curve; it is cut by the boundary segments into two pieces. Take one of the pieces

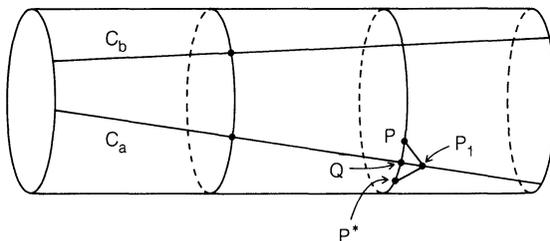


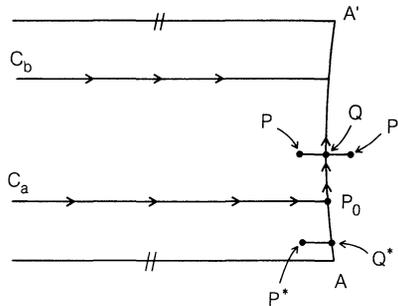
Fig. 7. We must show that the boundary components of  $R$ , namely  $\mathcal{C}_a$  and  $\mathcal{C}_b$  divide the ring domain into two identical rectangular strips of width  $\pi$

of the core curve; it is a geodesic, so its image under the map must also be a geodesic between the two endpoints. The only other such geodesic is the other segment of the core curve. Therefore these two segments are mapped into each other, and they must have equal lengths ( $= \pi$ ).

Let us finally show that the core curves and the boundary segments must be orthogonal to each other (this shows that the boundary segments cannot twist around the cylinder). Suppose they are not, as illustrated in Fig. 7. Consider the point  $P$  lying on the core curve, near the intersection point  $Q$ , and its image  $P^*$ . Clearly,  $l_{PQ} = l_{P^*Q}$ . Now consider a point  $P_1$  on the boundary segment, also near  $Q$ . The image of the segment  $PP_1$  must be the segment  $P^*P_1$ , which is of different length unless the core curve and the boundary segment are orthogonal. Since the core curves are the closed horizontal trajectories of the quadratic differential, the boundary segments are vertical trajectories of the quadratic differential. This argument is simply equivalent to saying that the map exchanging  $R$  and  $R^*$  is just a reflection about the boundary components. It applies as well to semi-infinite cylinders (corresponding to the punctured disks), thus we have shown that the boundary components of  $R$  cut every ring domain of the quadratic differential on  $\hat{R}$  into two identical flat rectangular strips of height  $\pi$ .  $\square$

(iii) Assume a boundary component  $\mathcal{C}_a$  of  $R$  appearing in some ring domain hits the boundary of its ring domain at a point  $P_0$  (the boundary of the ring domain is by definition a critical trajectory). We want to show that the boundary component  $\mathcal{C}_a$  cannot turn and become immediately part of the boundary of the ring domain.

Consider first the case when the point  $P_0$  is an ordinary point on the critical trajectory. Suppose that  $\mathcal{C}_a$  actually hits  $P_0$  and becomes part of the boundary, as illustrated in Fig. 8, where we show a ring domain whose boundary is the

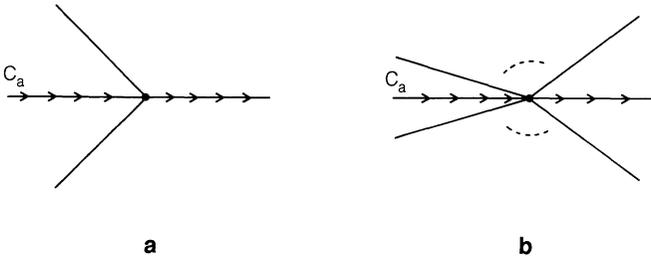


**Fig. 8.** It will be shown that the curve  $\mathcal{C}_a$ , representing a boundary component of  $R$  cannot hit the ordinary point  $P_0$  on the critical trajectory  $AA'$  and become immediately part of the critical trajectory. If this would happen  $P_0$  would become a first order pole

closed critical trajectory  $AA'$  ( $A$  and  $A'$  are the same point). Since this is a critical trajectory some ring domain is glued to the right of the line  $AA'$ . Consider the point  $Q$  on the line  $AA'$  and a point  $P$  nearby. The image of  $P$  under  $\mathcal{C}_a$  is  $P^*$ . Since  $\mathcal{C}_a$  also goes through  $Q$ , and the map is locally a reflection,  $P$  also maps into  $P'$ . Since the map is one to one  $P^* = P'$  and the same ring domain is also sitting to the right of  $AA'$ . Consider now the segment  $PQ$ , it has as image

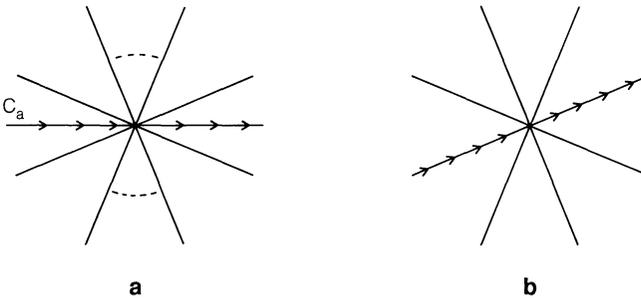
both the segment  $P^*Q^*$  and  $P'Q$ , but these two are clearly the same segment, thus  $Q^* = Q$ . This implies that the segment  $P_0Q$  is actually identified (glued) to the segment  $P_0Q^*$ , with  $P_0$  being a singular point. The point  $P_0$  becomes the center of a neighborhood of total angle  $\pi$ , thus a first order pole of the quadratic differential. This is a contradiction since the point  $P_0$  was supposed to be a regular point. At any rate the quadratic differentials in  $\Phi_{\hat{R}}$  have no first order poles. (If such pole were there, there would exist closed curves homotopic to the core curve of the ring domain and shorter than  $2\pi$ , such as the closed curve formed by  $QA'$  together with  $AQ^*$ .)

Let us now consider the possibility that  $P_0$  is a critical point, namely, an  $n^{\text{th}}$  order zero of the quadratic differential ( $n \geq 1$ ). An  $n^{\text{th}}$  order zero corresponds to a singularity with  $n+2$  prongs, where  $n+2$  ring domains join. Suppose we have a first order zero, as shown in Fig. 9. The boundary  $\mathcal{C}_a$  hits the zero symmetrically (since



**Fig. 9a, b.** If the boundary component of  $R$ , denoted as  $\mathcal{C}_a$  would hit an  $n^{\text{th}}$  order zero, with  $n$  odd, it must exit along a critical trajectory. **a** The case  $n = 3$ . **b** The general case

it is a vertical trajectory in the ring) and in order to preserve reflection symmetry of the metric around the singularity, it must exit along a critical trajectory, the one that is not on the boundary of the original ring domain. This is generic for all odd  $n$ , as shown to the right. This takes care of the case when  $n$  is odd. If  $n$  is even, symmetry around the zero requires that  $\mathcal{C}_a$  hit the zero and leave it along a vertical trajectory, as shown in part (a) of Fig. 10. This shows that  $\mathcal{C}_a$  cannot turn into a critical trajectory at all for even  $n$ . Thus we have shown that if  $\mathcal{C}_a$  hits a critical point  $P_0$ , it cannot turn immediately into a critical trajectory in the boundary of the corresponding ring domain. Since we dealt above with the case



**Fig. 10a, b.** If the boundary component of  $R$ , denoted as  $\mathcal{C}_a$  would hit an  $n^{\text{th}}$  order zero, with  $n$  even, it can happen in two ways. **a** it enters and exits along vertical trajectories, or **b** it enters and exits along a horizontal critical trajectory

when  $P_0$  is regular, this concludes the proof of (iii). Our proof does not rule out the possibility that  $\mathcal{C}_a$  hits an odd-order zero and turns into a critical trajectory not lying on the boundary of the original ring domain.  $\square$

We have established that the boundary components of  $R$  show up as vertical trajectories on each ring domain. Our result also implies that the reflection map acting on  $\hat{R}$  does not move points from one ring domain to another. The vertical trajectories representing boundary components cut the cylinders into two strips that are mapped into each other by the reflection map. The boundary components of  $R$  could also appear in the critical trajectories of the quadratic differential. Since critical trajectories, by definition, are horizontal trajectories, it follows that the quadratic differential is real all along the boundary components of  $R$ . This has a direct consequence.

**Corollary 1.** *The quadratic differential  $\varphi_{\hat{R}}$  arising from a mirror symmetric admissible set of curves in  $\hat{R}$  is mirror symmetric.*

*Proof.* The quadratic differential  $\varphi_{\hat{R}}$  induces a quadratic differential on  $R$  and one on  $R^*$ . We denote this as  $\varphi_{\hat{R}} = (\varphi_R, \varphi_{R^*})$ . Since  $\varphi_R$  is real along the boundaries of  $R$  it can be extended to the mirror surface to form a quadratic differential on the double surface [St, p. 19]. Thus we can construct another quadratic differential  $\varphi'_{\hat{R}}$  on  $\hat{R}$  using  $\varphi_R$  in  $R$ , and assigning to  $R^*$  the mirror image  $\varphi_R^*$ . We thus have  $\varphi'_{\hat{R}} = (\varphi_R, \varphi_R^*)$ . By definition,  $\varphi'_{\hat{R}}$  is a mirror symmetric quadratic differential. The quadratic differentials  $\varphi_{\hat{R}}$  and  $\varphi'_{\hat{R}}$ , are both defined in  $\hat{R}$ , and are equal to each other in some finite region on  $\hat{R}$ , namely, all over  $R$ . By analytic continuation, they must be identical all over  $\hat{R}$ . Therefore  $\varphi_{\hat{R}} = \varphi'_{\hat{R}}$ , and this shows that  $\varphi_{\hat{R}}$  is mirror symmetric.

## 5. A Generalized Extremal Problem

We have defined and studied in the previous section an interesting class of quadratic differentials, those based on mirror symmetric admissible sets of curves. Given a Riemann surface  $R$  there are an infinite number of mirror symmetric admissible sets of curves. Each one can, a priori, yield a different quadratic differential. We need a criterion to choose a particular quadratic differential out of the set. As we did for the case of classical closed string diagrams [Zw1], we will choose the quadratic differential with largest (reduced) norm, or (reduced) area. In this section we will prove some properties of this quadratic differential. In Sect. 6 we will establish that the quadratic differential of largest norm is unique and indeed it is open string quadratic differential we are after.

*Generalized Extremal Problem.* Consider a Riemann surface  $R$  with  $b > 0$  boundary components and with  $m \geq 0$  punctures at the boundaries (except for the disk with  $m \leq 1$  punctures). Denote by  $R^*$  the mirror image of  $R$  and by  $\hat{R}$  the double of  $R$ . Consider the set  $\Gamma_{\hat{R}}$  whose elements are all possible mirror symmetric admissible sets of curves  $\hat{\gamma}$  in  $\hat{R}$  (an infinite number). For each admissible set of curves in  $\Gamma_{\hat{R}}$  there is an associated mirror symmetric quadratic differential  $\varphi_{\hat{R}}(\hat{\gamma})$ . The collection of quadratic differentials obtained in this way is called  $\Phi_{\hat{R}}$ . We are interested in studying the quadratic differential(s) of largest norm in  $\Phi_{\hat{R}}$ .

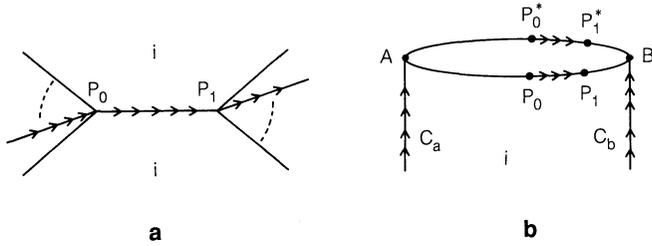
The above extremal problem is a mini-max problem (as explained to me by K. Strebel). One first minimizes the area for a given choice of curves, and then maximizes over the possible choices of curves. The only fact one needs to show at this stage is that at least one of the quadratic differentials in  $\Phi_{\hat{R}}$  attains the maximal norm. The proof given in [Zw1], where we discussed the same issue for closed Riemann surfaces, applies without modification for the present case. Thus, there is at least one quadratic differential of maximal norm. In fact, we will show in Sect. 6 that this quadratic differential is unique. The rest of the present section will be devoted to establishing some properties of the quadratic differential of maximal norm. Denote by  $\hat{\phi}$  any such quadratic differential.

**Lemma 3.** *The quadratic differential  $\hat{\phi}$  induces on  $R$  a quadratic differential such that any nontrivial open Jordan curve homotopic to a critical path is longer or equal to  $\pi$ .*

*Proof.* Denote the open curve by  $\gamma_0$  and assume it is shorter than  $\pi$ . We will show that this contradicts the assumption that we are dealing with the quadratic differential of maximal norm  $\hat{\phi}$ . The double  $\hat{\gamma}_0$ , made up by joining  $\gamma_0$  to its mirror image  $\gamma_0^*$ , is twice as long as  $\gamma_0$ , and is therefore smaller than  $2\pi$ . The curve  $\hat{\gamma}_0$  is a nontrivial Jordan closed curve [Zw3, Lemma 6]. Thus  $\hat{\gamma}_0$  is a curve that can belong in a mirror symmetric admissible set. The curve  $\hat{\gamma}_0$  cannot be homotopic to any of the curves in the admissible set for the quadratic differential  $\hat{\phi}$ , since if it were, it would have to be longer than  $2\pi$ . The admissible set associated to  $\hat{\phi}$  can be taken to contain a representative curve around each puncture, chosen far away from any critical graph, and a representative curve for each annulus. This latter curve can be taken to be a core curve in the middle of each annulus. The curve  $\hat{\gamma}_0$  may intersect the representatives of the admissible set. Nevertheless, since it is homotopic to a closed critical trajectory, one can choose a representative  $\bar{\gamma}_0$  in that homotopy class that does not intersect the other representatives in the admissible set for  $\hat{\phi}$ . This can always be done since the annuli have finite heights and one can push the curve close enough to the critical trajectory and away from the core curve in the middle of the annuli. (The representative cannot in general be chosen to be the critical closed path since this may not be a Jordan curve.) We have shown that our representative does not intersect nor it is homotopic to any of the curves in the admissible set. Therefore it can be added to the present admissible set  $\tilde{\gamma}'$  giving a new admissible set  $\tilde{\gamma}'$ . Since the closed curve  $\hat{\gamma}_0$  is shorter than  $2\pi$  it means that the present quadratic differential does not minimize the area under the condition that all curves in the homotopy classes of  $\tilde{\gamma}'$  be greater than  $2\pi$ . The minimum area metric  $\varrho'$  for the set  $\tilde{\gamma}'$  must be different from the minimal area metric  $\varrho$  for the set  $\tilde{\gamma}$ , and its area must therefore be larger (since the metric  $\varrho'$  is also admissible for the set  $\tilde{\gamma}$ , and is different from the minimal area metric  $\varrho$ ). Thus the original quadratic differential cannot be the one of largest norm in  $\Phi_{\hat{R}}$ . This is the contradiction.  $\square$

**Lemma 4.** *No part of the boundary of  $R$  can lie on a critical trajectory of  $\hat{\phi}$ .*

*Proof.* Assume part of the boundary of  $R$  lies along a critical trajectory as shown in Fig. 11. The part of the boundary in question must extend between two zeros  $P_0$  and  $P_1$ , since otherwise it would violate Theorem 2, Part (iii) Note again, that since the reflection map does not move points into different ring domains, we must have the same ring domain, denoted as “ $i$ ,” to the two sides of the  $P_0P_1$  critical trajectory. The segment  $P_0P_1$  cannot include the two points  $A$  and  $B$



**Fig. 11a, b.** It will be shown that a boundary component of  $R$  cannot lie on a critical trajectory of the quadratic differential of maximal norm. **a** We assume it does, namely a boundary component extends along the critical segment  $P_0P_1$ . The same ring domain, denoted as  $i$  must appear to each side of  $P_0P_1$ . **b** A picture of the ring domain  $i$ , with the critical segment and its own boundary components  $C_a$  and  $C_b$

where the vertical trajectories  $C_a$  and  $C_b$  of ring “ $i$ ” (representing boundaries of  $R$ ), hit the boundary of the ring domain Fig. 11b). If the segment included  $A$  or  $B$  in its interior, we would have crossing of boundary components of  $R$ , which is impossible. If the endpoints of the segment would correspond to either  $A$  or  $B$  we would violate Theorem 2, Part (iii). Finally note that beyond the zeroes  $P_0$  and  $P_1$ , the boundary component of  $R$  cannot remain in the boundary of our ring domain; it must exit the ring domain at those points. In other words, there are no more zeroes in the segment.

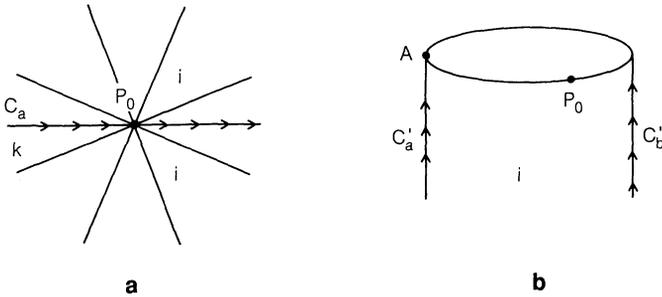
The pattern of identification in the ring domain is shown to the right. The segment  $P_0A$  is an open segment with endpoints on boundaries of  $R$ , lying on a critical trajectory. It is clear that near enough to  $P_0$  and to  $A$  the segment does not include any piece of the boundary of  $R$ . Assume for the time being that the segment  $P_0A$  is an open Jordan curve (namely, does not include in its interior any point of the boundary of  $R$ ). This open segment must be nontrivial because its double, a closed Jordan critical path, is nontrivial [Zw3, Sect. 2,3], and if the double of an open curve is nontrivial, then the open curve is nontrivial [Zw3, Lemma 6]. It follows from Lemma 3 that the length  $l_{P_0A}$  of the segment must exceed  $\pi$ . This is in direct contradiction with Fig. 11b, that shows that the length  $l_{P_0A}$  is less than  $\pi$  (because  $A$  must be equidistant from  $P_0$  and  $P_0^*$ , and  $l_{P_0P_1} \neq 0$ ). This proves Lemma 4 up to the assumption with which we deal now.

We assumed that the segment  $P_0A$  is a simple Jordan open curve with endpoints on boundary components of  $R$ . This may not be true, but, as we show next, it is always possible to find a new segment that will play the desired role. Several things can happen inside the segment  $P_0A$ : it may happen that both points  $A$  and  $P_0$  repeat themselves; there can be self identification of subsegments within  $P_0A$ ; or some other segments of the boundary of  $R$  may lie within  $P_0A$ . Begin at  $A$  and go towards  $P_0$ . Denote by  $B$  the first point when we encounter a boundary component. This point may just be the original point  $P_0$  or may be a point we find earlier. From now on we concentrate on the open segment  $AB$  which does not contain in its interior any point belonging to the boundary of  $R$  (if  $B$  happened to be a copy of  $A$ , the segment would be closed). There may be self identifications inside the  $AB$  segment, or some zero may repeat several times, any of these things makes the segment  $AB$  a curve with self-intersections. If a segment lying between zeroes  $Z_0$  and  $Z_1$  is self identified, then the sequence of zeroes from  $A$  to  $B$  must read  $A \dots Z_0Z_1 \dots Z_1Z_0 \dots B$ . From the segment  $AB$

we find a segment without self-intersections by beginning at  $A$  and each time we find a letter that is repeated later we jump up to its latest occurrence in the series. We continue in this way until we reach point  $B$ . This new segment contains no boundary points except  $A$  and  $B$  and no self-intersections. It is shorter than  $\pi$  because it is shorter than  $P_0A$ . This is the desired open Jordan critical segment. The only remaining complication could arise, as mentioned above, if  $B$  is just a copy of point  $A$ . This would make the segment, which is already free of self intersections or boundary points in its interior, into a closed Jordan critical segment, thus nontrivial. It is possible, however, to deform the segment infinitesimally in order to make it into an open segment, without increasing its length beyond a specified amount that can be chosen at will. Near the beginning point of the curve we just deform it such that the curve begins on  $\mathcal{C}_a$  a bit below  $A$  and joins into the critical trajectory immediately. The double of this open curve would be a nontrivial closed curve since it is homotopic to a closed curve that is nontrivial (as analyzed in [Zw3, Sect. 2.3, Fig. 1]). Thus the open curve in question is nontrivial, and being shorter than  $\pi$  this gives the desired contradiction.  $\square$

**Lemma 5.** *The boundary of  $R$  cannot go through a zero of  $\hat{\phi}$ .*

*Proof.* It is clear from Fig. 9 that this cannot happen for odd  $n$ . In those cases, part of the boundary of  $R$  must lie on a critical trajectory, and this was ruled out in Lemma 4. The only remaining possibility is that sketched in Fig. 10a dealing with even  $n$ . Not even this is possible, as we show next. Consider Fig. 12 where



**Fig. 12a, b.** It will be shown that the boundary component  $\mathcal{C}_a$  of  $R$  cannot hit a zero of the quadratic differential of maximal norm. **a** Note the ring domain  $i$  that must appear symmetrically. **b** The ring domain  $i$ , where we see its own boundary components  $\mathcal{C}'_a$  and  $\mathcal{C}'_b$

the piece  $\mathcal{C}_a$  of the boundary of  $R$  is denoted by the line with arrows that hits the zero  $P_0$  from the  $k^{\text{th}}$  ring domain. The  $i^{\text{th}}$  ring domain, due to the reflection symmetry, must appear two times as shown. Assume the  $i^{\text{th}}$  ring is not the same ring as the  $k^{\text{th}}$  ring. The  $i^{\text{th}}$  ring domain is shown to the right. Point  $A$  (where  $\mathcal{C}'_a$  hits the boundary of the  $i^{\text{th}}$  ring) cannot coincide with  $P_0$ , nor can one be a copy of the other. If this would be the case, we would have in Fig. 12a another boundary component of  $R$  hitting the zero, and this cannot happen since it would imply intersection of boundary components of  $R$ . It follows, just as before, that the critical segment  $P_0A$  with boundary endpoints must be shorter than  $\pi$ , and this is impossible, as discussed in Lemma 4. If the  $k^{\text{th}}$  and  $i^{\text{th}}$  ring domains are actually the same, then Fig. 12b still applies, it describes the  $i = k$  domain, except that now point  $A$  is a copy of  $P_0$ , and  $\mathcal{C}'_a = \mathcal{C}_a$ . We can now apply the arguments

in Lemma 4 to the closed path  $AP_0$  (namely, modify it infinitesimally to make it open, and show it is shorter than  $\pi$ , and nontrivial).  $\square$

## 6. On the existence of Open String QDS

We have done most of the work necessary in order to show that the generalized extremal problem gives us an open string diagram. We have learned so far that the surface  $R$  with the metric arising from the quadratic differential of maximal area is built by joining together strips of height  $\pi$ . We need as many semi-infinite strips as there are punctures, and as many finite strips as there are ring domains in  $\hat{R}$ . We also know that the strips must be joined across the horizontal edges, and that the vertical edges correspond to the boundaries of  $R$ . We also proved that after the gluing of the horizontal edges, the corners of the strips are regular points. This is so because the boundaries of  $R$  do not hit any zero of the quadratic differential. Finally, we know that *every open critical Jordan path on  $R$  must be longer than  $\pi$* . This is so because these paths are nontrivial, and thus Lemma 3 applies.

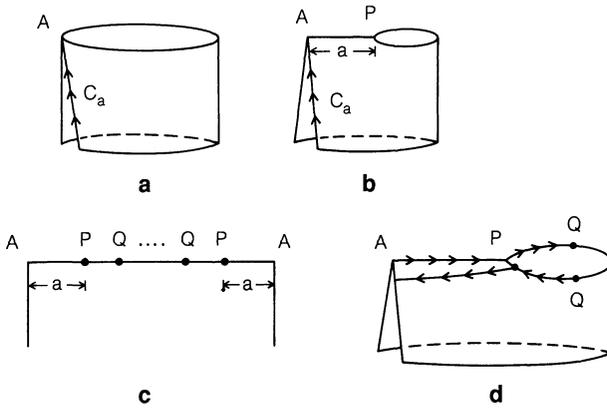
Our next step will be to analyze possible identification structures on a single strip. The result we will establish next is given in the following lemma.

**Lemma 6.** *Consider a horizontal edge of a strip in a quadratic differential of maximal norm. The two endpoints of the horizontal edge cannot be identified. No zero can appear on the edge more than once.*

*Proof.* Consider a strip whose horizontal boundary, of length  $\pi$  is denoted as  $AA'$ . Recall that points  $A$  and  $A'$  belong to the boundary of  $R$ . No other point in the horizontal edge can belong to the boundary of  $R$ , as a consequence of Lemmas 4 and 5.

Let us assume for awhile that points  $A$  and  $A'$ , are not identified by the gluing that takes place at the horizontal edge of the strip. Under this assumption, we now show that any zero of the quadratic differential cannot show up more than once on the line  $AA'$ . In particular that implies that there cannot be self identification of segments along the  $AA'$  line. This follows because self identification of a segment requires that two zeroes appear twice. Suppose we have a sequence of zeroes in which at least one zero repeats itself (the sequence must be finite). We will extract a path beginning at  $A$  and ending at  $A'$  which is a Jordan open critical path, and is shorter than  $\pi$ . The existence of such open curve is a contradiction. The open path begins at  $A$  and stops at the first zero  $Z_0$  that is repeated in the sequence, we then jump to the last place where  $Z_0$  appears in the sequence, and the open path continues its way down there. This is repeated for the left-over sequence, and repeated until we end in point  $A'$ . This gives us the desired open Jordan critical curve. Since we assumed that at least some zero repeats itself, we must have jumped some finite length and therefore the resulting segment is shorter than  $\pi$ , which was the length of  $AA'$ . Thus we have the desired contradiction.

We wish to rule out the possibility that  $A$  and  $A'$  be identified. If points  $A$  and  $A'$  are identified in the way shown in Fig. 13a, this is problematic. Once this strip is glued to its double, the neighborhood of  $A$  would already have a full  $2\pi$  angle. When the horizontal edge near point  $A$  is sewn with other domains, the total angle at  $A$  would become bigger than  $2\pi$ . Point  $A$  would therefore be

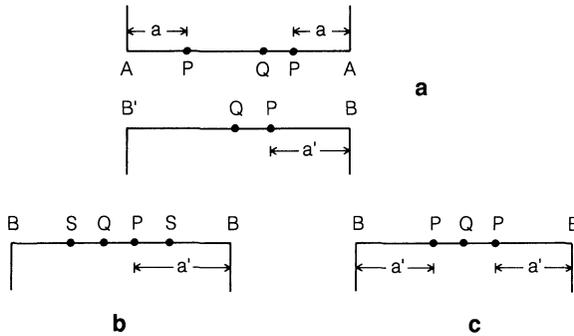


**Fig. 13a-d.** Analyzing if the endpoints of a critical horizontal segment on a strip can be identified. **a** Identification of the endpoints only is ruled out. **b** If at all possible, the endpoints are identified as part of the segment  $AP$ . **c** The critical trajectory with such identification. **d** A Jordan open path homotopic to the critical trajectory

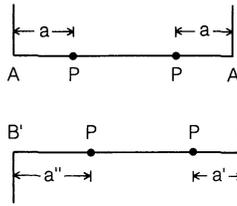
a zero of the quadratic differential. But this is impossible, since we showed that the boundary of  $R$  does not hit zeroes of the quadratic differential. Therefore, points  $A$  and  $A'$  must be identified as part of a segment of identification in the critical trajectory as shown in Fig. 13b, c. Let  $a$  be the length of the segment  $AP$ . We have shown that  $a > 0$ . Moreover  $a < \pi/2$ , since for  $a = \pi/2$  we would have a degenerate Riemann surface.

Let us now show that in this configuration no other zero in the sequence  $P \dots P$  can be repeated (not even  $P$  may appear once more). If some zero is repeated we jump it until we get to its last copy. In this way we get an open path beginning at  $A$  going to  $P$ , and then skipping its way back to  $P$  and then to  $A$ . It would be shorter than  $\pi$ . This path retraces the segment  $AP$ , but it can be deformed infinitesimally without increasing its length to the value  $\pi$ , and making it a simple open Jordan curve homotopic to a critical open trajectory (see Fig. 13d). This curve is nontrivial since its double is a nontrivial closed curve (as discussed in [Zw3, Sect. 2.3]). The existence of this curve is in contradiction with Lemma 3.

Now we can show that it is actually impossible to have  $A = A'$ . Suppose we have the arrangement of Fig. 14, where the sequence of zeroes goes like  $PQ \dots P$ , the first zero after  $P$  being  $Q$ , and with the bottom and top segments  $AP$  glued to each other. Since self-identifications on a strip are impossible, the segment  $PQ$  must appear in some other strip, shown below. In this strip there may be additional zeroes (not shown). The endpoints  $B$  and  $B'$  cannot be identified with  $A$  since the neighborhood of  $A$  is complete. Assume  $B$  and  $B'$  are not identified. The segment  $APB$  is an open critical Jordan curve since  $A \neq B$  and it has no self intersections. The condition that it be longer or equal to  $\pi$  gives us  $a + a' \geq \pi$ . Moreover the curve  $APQB'$  is another open critical Jordan curve, giving us  $l_{AP} + l_{PQB'} \geq \pi$ , which yields  $a \geq a'$ . The two inequalities imply that  $a \geq \pi/2$ , which in light of Fig. 14a requires  $a = \pi/2$ . But this is a contradiction since it requires the surface to be singular. This argument needs slight modification for the case when  $B$  is identified with  $B'$ . There are two possibilities, as shown in Fig. 14b, c. In (b) the identification segment is  $BS$  and the above arguments hold in exactly the same way. In (c) the identification



**Fig. 14.** a A strip with endpoints identified with at least one zero  $Q$  in the  $PP$  segment. The segment  $QP$  must be identified with a similar segment in another strip, denoted as  $B'B$ . The necessary modifications when  $B$  and  $B'$  are identified are dealt with in **b** and **c**



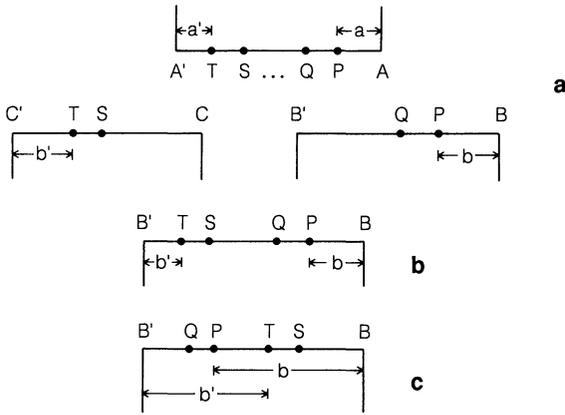
**Fig. 15.** A strip with identified endpoints and no zero in the segment  $PP$ . This segment is identified with a similar segment in another strip ( $BB'$ ). Actually  $B$  and  $B'$  must be identified. The final conclusion is that endpoints of the critical horizontal of a strip cannot be identified

segment is  $BP$ . Here the path  $APB$  is still an open Jordan critical curve and its length condition gives  $a+a' \geq \pi$ , that implies that either  $a$  or  $a'$  is greater or equal to  $\pi/2$ . Again, this makes the surface singular, and is therefore impossible. So we have shown that the identification of segments  $AP$  in Fig. 14 is not consistent if we have at least one zero ( $Q$ ) in the  $PP$  closed curve.

Let us now address the case when there is no zero in the closed curve  $PP$ , as shown in Fig. 15. The full closed  $PP$  segment must be identified with another  $PP$  segment in another strip, as shown below. Since the point  $P$  appears at least twice in this strip, this is only possible if points  $B$  and  $B'$  are identified. In fact they must be identified via the segment  $BP$ , since if it would be via another shorter segment, then  $P$  appearing more than once would be a contradiction. It then follows that  $a' = a'' = a$ . The open critical Jordan path  $APB$  requires that  $a + a' = 2a \geq \pi$ , where from we deduce that  $a = \pi/2$ , which again is inconsistent. This shows that having  $A = A'$ , and no zeroes in the closed curve  $PP$  is impossible, and together with the discussion of the above paragraph, that  $A$  and  $A'$  cannot be identified at all.  $\square$

We will now use the above information to show in the next lemma that the horizontal boundaries of the strips in the quadratic differential of maximal norm are glued in a very simple way.

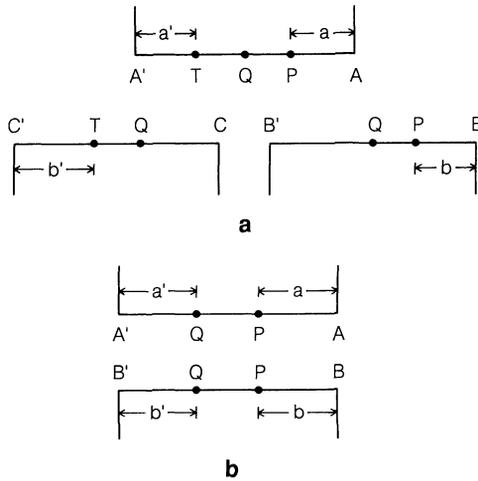
**Lemma 7.** *In a quadratic differential of maximal norm the zeroes can only appear in the midpoint of the horizontal edges of the strips.*



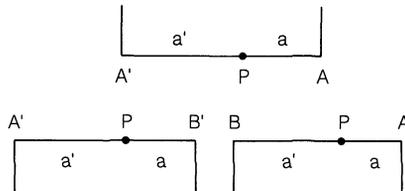
**Fig. 16a–c.** Studying patterns of gluing. **a** We assume there are four or more zeroes in the horizontal trajectory  $A'A$ . The segments  $TS$  and  $QP$  appear in other strips  $C'C$  and  $B'B$  respectively. If the segments  $TS$  and  $QP$  appear in the same strip ( $B'B$ ) there are two cases to consider, as shown in **b** and **c**. The final conclusion is that  $a = a' = \pi/2$  which implies that all the zeroes collapse to the midpoint of the horizontal segment

*Proof.* Let us now address the general configuration shown in Fig. 16. In part **(a)** we have a strip with horizontal edge  $AA'$  and let us assume there are several zeroes in this segment:  $P, Q, \dots, S, T$ . The zero  $P$  is the first zero we meet as we go down the line from  $A$ , and  $T$  is the last zero in this segment. The segments  $AP$  and  $TA'$  of lengths  $a$  and  $a'$  respectively, must be glued to other strips. The segment  $PQ$  is glued to the strip  $BB'$ , shown below, and the segment  $ST$  is glued to the strip  $CC'$  (the case when  $PQ$  and  $ST$  are glued to the same strip is considered later). Let us show points  $A$  and  $B$  are not identified. If they were, it can only happen if point  $P$  would be the first zero from  $B$ , and the segment  $AP$  would be identified with the segment  $BP$ . This is impossible, however, since this identification would be followed directly by the identification of segment  $PQ$ , and as a consequence  $P$  could not be a zero. Therefore  $APB$  is an open Jordan critical curve, and we must have  $a + b \geq \pi$ . Moreover,  $A$  and  $B'$  cannot be identified, since  $P$  would have to appear more than once. Thus the open Jordan critical curve  $APQB'$  gives the condition  $a \geq b$ . These two inequalities imply that  $a \geq \pi/2$ . In an exactly analogous fashion the nontrivial open Jordan critical curves  $A'TC'$  and  $A'TSC$  imply that  $a' \geq \pi/2$ . Since both  $a$  and  $a'$  appear in the same strip the only possibility is that  $a = a' = \pi/2$ . This being the case, all the zeroes ( $P, Q, \dots, S, T$ ) must collapse into a single one at the middle of the strip. Thus the strips must be glued together with zeroes that can only appear at the middle of the strip. This conclusion is not changed if both segments  $PQ$  and  $ST$  are glued to the same strip  $BB'$ , as shown in Fig. 16c, d. All of the above arguments hold with the lengths  $b$  and  $b'$  defined as shown.

Strictly speaking, the above argument shows that if there are four or more zeroes in the line  $AA'$ , they must collapse in the middle. One must repeat the argument for the case when there are three zeroes, or two zeroes in the line  $AA'$ . Since the arguments are almost identical to that given above, we just show the relevant configurations in Fig. 17a, b. For the case of three zeroes,  $P, Q$ , and  $T$ , the two segments  $PQ$  and  $QT$  cannot appear in the same strip again since  $Q$  cannot appear twice, and therefore they would have to appear one after the other, as in the original strip, making  $Q$  into a regular point.



**Fig. 17a, b.** Analyzing patterns of gluing when there are three zeroes in the horizontal trajectory **a**, and when there are two zeroes **b**. Again, all zeroes must collapse to the midpoint of the horizontal segment



**Fig. 18.** A strip horizontal segment has a single zero. A short argument shows that  $a = a' = \pi/2$ . Thus the gluing pattern is that of open string theory

We can now finally address the simplest configuration, that having a single zero on the horizontal boundary, as shown in Fig. 18. Here we show the horizontal edge  $AA'$  with the zero at  $P$ . Let the length of the segment  $AP$  be  $a$ , and that of the segment  $A'P$  be  $a'$ . The segment  $AP$  is glued to a segment in the horizontal edge  $AB$  shown below. The segment  $A'P$  must be glued to a segment in yet another horizontal edge  $A'B'$ . The open Jordan critical curves  $APB'$  and  $A'PB$  yield the conditions  $a \geq \pi/2$  and  $a' \geq \pi/2$ , respectively. These conditions imply  $a = a' = \pi/2$ , and therefore we have shown that the zero must appear at the midpoint of the vertical segment.  $\square$

**Theorem 3.** *The quadratic differential of maximal norm in  $\Phi_{\hat{R}}$  is unique and induces on  $R$  an open string quadratic differential.*

*Proof.* It has been established that the quadratic differential of maximal norm has as characteristic ring domains in  $R$  strips with horizontal trajectories of length  $\pi$ , whose horizontal edges are critical trajectories, and whose vertical boundaries correspond to boundaries of  $R$  and are vertical trajectories. We have shown that the endpoints of a horizontal edge cannot be identified, and that all zeroes of the quadratic differential lie on the midpoints of the horizontal edges. Thus, all of

our work has established that any quadratic differential of maximal norm gives us an open string quadratic differential. This establishes the existence, for any  $R$ , of an open string quadratic differential.

Assume  $\phi$  and  $\phi'$  are two different open string quadratic differentials. As we have shown in Sect. 2, an open string quadratic differential gives rise to a minimal area metric under the condition that all nontrivial Jordan open curves be longer or equal to  $\pi$ . This metric must be unique, thus  $\phi$  and  $\phi'$  must give rise to the same metric. Because of Lemma 2 (Sect. 2), the quadratic differentials must be the same. This concludes our proof of uniqueness of the quadratic differential of maximal norm in  $\Phi_{\hat{R}}$ .  $\square$

Theorem 3 together with Theorem 1 are summarized by Theorem 4 in the introduction. As discussed in Sect. 2 the uniqueness and existence of open string quadratic differentials, given the isomorphism between the set of Feynman diagrams and the set of string diagrams, implies the single cover or moduli space, as given in Theorem 5.

## References

- [BS] Bluhm, R., Samuel, S.: The off-shell Koba-Nielsen formula. Nucl. Phys. B **323**, 337 (1989)
- [Gi] Giddings, S.: Conformal techniques in string theory and string eidl theory. Phys. Rep. **170**, 167 (1988)
- [GMW] Giddings, S., Martinec, E., Witten, E.: Modular invariance in string field theory. Phys. Lett. **176 B**, 362 (1986)
- [GW] Giddings, S., Wolpert, S.: A triangulation of moduli space from light-cone string theory. Commun. Math. Phys. **109**, 177 (1987)
- [KKS] Kugo, T., Kunitomo, H., Suehiro, K.: Non-polynomial closed string field theory. Phys. Lett. **226 B**, 48 (1989)
- [KS] Kugo, T., Suehiro, K.: Nonpolynomial closed string field theory: action and gauge invariance. Nucl. Phys. B **337**, 434 (1990)
- [Og] Ogura, W.: Combinatorics of strings and equivalence of Witten's string field theory to Polyakov's string theory. Progr. Theor. Phys. **79**, 936 (1988)
- [Sa] Samuel, S.: Solving the open bosonic string in perturbation theory. Nucl. Phys. B **341**, 513 (1990)
- [SaZw] Saadi, M., Zwiebach, B.: Closed string field theory from polyhedra. Ann. Phys. **192**, 213 (1989)
- [St] Strebel, K.: Quadratic differentials. Berlin, Heidelberg, New York: Springer 1984
- [Th] Thorn, C.: String field theory. Phys. Rep. **174**, 1 (1989)
- [Wi] Witten, E.: Noncommutative geometry and string field theory. Nucl. Phys. B **268**, 253 (1986)
- [Zw1] Zwiebach, B.: How covariant closed string theory solves a minimal area problem. Commun. Math. Phys. **136**, 83 (1991); Consistency of closed string polyhedra from minimal area. Phys. Lett. B **241**, 343 (1990)
- [Zw2] Zwiebach, B.: Quantum closed strings from minimal area. Mod. Phys. Lett. A **2**, (32) 2753 (1990)
- [Zw3] Zwiebach, B.: Minimal area problems and quantum open strings. Commun. Math. Phys. (in press)
- [Zw4] Zwiebach, B.: Quantum open string theory with manifest closed string factorization. Phys. Lett. B **256**, 22 (1991); The covariant open-closed string theory I, II. MIT preprints, MIT-CTP-1909, 1910 (to appear)
- [Zw5] Zwiebach, B.: Interpolating open-closed string field theories. MIT preprint MIT-CTP-1911 (to appear)