APPEARED IN BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 29, Number 1, July 1993, Pages 130-136

BOOK REVIEW

Combinatorics of train tracks, by R. C. Penner with J. L. Harer. Annals of Math. Stud., vol. 125, Princeton Univ. Press, Princeton, NJ, 1992, xii+216 pp., $19.95. ISBN 0-691-02531-2

The theory of surfaces or Riemann surface theory, from its inception, has often been driven by computational and/or combinatorial questions. The roots of the theory lie in the study of algebraic integrals in the complex plane. Since that time, a significant part of surface theory has been devoted to the pursuit of combinatorial schemes to understand better some of the deeper structure, be it topological, geometric, algebraic, or analytic. The early work of Fagnano and later Euler, Abel, and Jacobi focused on the addition theorems for abelian integrals (see, e.g., Siegel [12]). In a different linguistic setting, we find this stream of ideas a still active part—and a spiritual foundation—of modern algebraic geometry. It is implicit in the constructions of Riemann, made concrete by Hurwitz, that a Riemann surface as a topological object may be defined by combinatorial data describing how to glue polygons together.

The complex analytic as well as the algebraic geometric approach to Riemann surface theory over most of the past one hundred years has been quite a bit less explicit. Emphasis was placed on broad general properties while interest in explicit algorithms lagged. Those of us in the complex analytic and topological wings of the Riemann surface community (as contrasted with those who study the same object under other names and guises) were treated to two jolts in the past two decades. The first came from Bill Thurston in the mid-1970s and underlies much of the substance of this review, while the second started about 1980 and came from particle physicists. The physicists actually wanted to do numerical computations, such as integration, in spaces of Riemann surfaces—a task for which we were quite unprepared.

The natural equivalence relation among Riemann surfaces is that of conformal or holomorphic equivalence. Not all topologically equivalent surfaces are conformally equivalent. The space of conformal equivalence classes of Riemann surfaces of a fixed finite topological type[1] is called the moduli space $\mathcal{M}$ and was first considered by Riemann, although its analytic structure was not given until about 1960 by Ahlfors and Bers (see, e.g., Nag [10]).

Isolated combinatorial approaches and methods for the study of Riemann surfaces had appeared intermittently over a period of about one hundred years. In the last two decades, Thurston created machines whose use enabled us to study the

[1]Here I am being sloppy by not distinguishing between punctures and bigger holes.

structure of surfaces in an organized fashion. The early Thurston methodology, in Bill Harvey's phrasing, consisted of taking a combinatorial situation, throwing it into a huge projective space to complete it, and then pulling back via nice finite sets of coordinates to a combinatorial situation. This concept is marvelous; it has spiritual roots in such topological constructs as classifying spaces. Thurston used the technique methodically and with awesome success. A major part of his success can be attributed to his ability to associate geometric objects to the ideal points added by the process of compactification.

I will trace some of the historical development of the ideas underlying the theory and applications of train tracks before discussing the modern treatments, of which the book under review is the most complete.

We will start in 1882, when Klein, in the process of "proving" the uniformization theorem, was starting to give coordinates to spaces of Riemann surfaces. His biggest difficulty and the only serious gap in his argument arose from the problem of deciding on an adequate description of a Riemann surface. By tradition he had to think of them as plane algebraic curves, but, actually, the objects he considered were finite volume hyperbolic surfaces or orbifolds. To avoid unnecessary complications, we will restrict our attention here, unless otherwise stated, to compact surfaces. He looked at a Riemann surface as a deformation of a discrete cocompact hyperbolic group. The space of such deformations is one of the many forms in which appears the Teichmüller space T_g of surfaces of genus g. Klein did not give a parametrization of the space. He simply thought of it as a collection of hyperbolic polygons with identifications or as the groups generated by the transformations which perform the identifications. Later Fricke (see Keen [7]) looked at some parameters, based on geometric dissections of surfaces to obtain global coordinates for the Teichmüller space. Unfortunately, the Teichmüller space is not the same as the moduli space of Riemann which we have previously mentioned. The moduli space may be obtained as a quotient of T_g by the mapping class group in genus g (see below). The projection is a branched covering.

We will need to consider a variant of Fricke's work. Suppose we are given a compact hyperbolic surface, that is, a compact surface S of genus $g > 1$ and a Riemannian metric of curvature -1 on S. S may be realized as the quotient, of the upper half plane model of the hyperbolic plane, by a group G of real möbius transformations. The group G is uniquely determined up to conjugation in the group of real möbius transformations. Suppose α is a simple closed homotopically nontrivial curve on S. Then in the free homotopy class of α there is a unique curve of minimal length which is again simple. So we might as well assume that α is this simple closed geodesic. If $\gamma \in G$ is a deck transformation covering the closed geodesic α, then the length l of α in the hyperbolic structure on S is related to the trace τ of γ by $\tau = \pm 2\cosh(l/2)$. It is a standard computation (see, e.g., Abikoff [1] among many others) that the lengths of sufficiently many simple closed curves completely determines the hyperbolic structure of S and, hence, also its conformal equivalence class.

Much later, Fenchel and Nielsen described the Teichmüller space using some lengths and angles (again, this may be found in many places; I know Abikoff [1] best). They parametrized the same space by gluing together triply connected domains or, in modern jargon, pants. (A *pants* is a closed disk from which we have removed the interiors of two disjoint closed subdisks.)

The use of pants in the study of surfaces had already been introduced by Dehn [3].

Dehn was one of the first mathematicians to introduce algebraic and combinatorial methods into topology. Among many other things, Dehn was interested in how an autohomeomorphism h acts on a surface S. Clearly, h permutes the (simple) closed curves on S. It also respects homotopy of curves, so, up to an indeterminacy caused by h not fixing the base point, h induces an automorphism $\pi_1 h$ of $\pi_1 S$. Any indeterminacy disappears when we consider the coset H, containing h, of the inner automorphisms of $\pi_1 S$ in the group of automorphisms of $\pi_1 S$. H is called the *mapping class* of h. It is well known—and apparently due to Dehn—that every outer automorphism of $\pi_1 S$ modulo inner ones is induced by a homeomorphism of S. A particularly nice set of generators of the mapping class group is given by the *Dehn twists*. Each of these homeomorphisms, later rediscovered by Lickorish, is the identity off a neighborhood of a simple closed curve. Unfortunately, the action of a general mapping class on either $\pi_1 S$ or T_g is rather difficult to see in terms of the Dehn twists, and this action is not given in terms of any natural coordinates. But Dehn did not stop there. He came up with another brilliant idea. He used a pants decomposition of S to give a combinatorial description of a simple closed curve α. This requires not only the choice of a pants decomposition but of a distinguished curve on each pants. One then counts how many times α intersects the distinguished curves and the border curves of the pants. From these counts one can recover the curve uniquely up to free homotopy. If you know a curve, you know its length in some hyperbolic metric. If you know the lengths of enough curves, you know the surface. Unfortunately, the map, associating to a closed curve the length of the geodesic in its free homotopy class, is not invertible; i.e., one cannot recover the homotopy class of curves from the unordered sequence of lengths of closed geodesics. In particular, we cannot determine mapping classes from lengths of curves. Both train tracks and the apparently more efficient methods of Mosher [9] can be used to give combinatorial descriptions of mapping classes.[2]

Dehn's description was rediscovered by Thurston in the mid-1970s. He made it into a theory by taking the numbers that are associated to a curve and using them as coordinates for a space of finite systems of simple closed curves. He was even able to compactify the space by adding ideal points which correspond to equivalence classes of (slightly singular) measured foliations (details can be found in [4]). I had two main difficulties in understanding Thurston's ideas. The first was how to recover a curve from the coordinates—that is not really difficult, at least visually. The main problem is how one changes coordinates—each choice of coordinates depends on a choice of pants decomposition and distinguished curve. In other words, the coordinate transition functions are complicated by the topological choices that one must make. The next observation was that this is a topological construct—it does not use any hyperbolic geometry. Perhaps iuniqueness of geodesics would start to ease the problems caused by too many topological choices.

The process of introducing geometry into the description of the simple closed curves has several consequences. After the introduction of geometry, the measured foliations were replaced as objects of study by measured geodesic laminations. These are completely unique once a hyperbolic structure is fixed on the surface but are really independent of any choice of hyperbolic structure. One can move back and forth between the descriptions afforded by foliations and laminations (see, e.g.,

[2] I am unaware of any direct comparisons between the two methods; however, Mosher's technique uses the available data more efficiently.

Levitt [8]).

Since train tracks, which ultimately are the subject of this review, are used to give local coordinates to the spaces of measured geodesic laminations with compact support, we need to define the latter concept. A *measured geodesic lamination with compact support* on a finite volume hyperbolic surface S is a pair $(\mathcal{L}, \mu)$ where $\mathcal{L}$ is a compact set given as the union of a collection of disjoint simple (closed or open) geodesics on S and μ is a measure on the family of smooth curves transverse to $\mathcal{L}$. $\mathcal{L}$ has a local product structure, and μ is required to respect this structure. In the following we assume all our laminations have compact support since our surfaces are assumed to be compact.

In the generic (uniquely ergodic) situation, $\mathcal{L}$ determines μ up to a scalar multiple, and it is common to refer to $\mathcal{L}$ as a measured geodesic lamination. If $\mathcal{L}$ consists of a finite number of simple closed geodesics, thenμ can just count the number of intersections of $\mathcal{L}$ with a curve meeting it transversely. The topology on the measured laminations is chosen so that these counting measures are dense even when $\mathcal{L}$ consists of a single simple closed geodesic.

An alternate description of a neighborhood of a measured lamination $(\mathcal{L}, \mu)$ consists of the measured laminations $(\mathcal{L}', \mu')$ for which $\mathcal{L}'$ is near $\mathcal{L}$ and μ' is a measure or weight, in a sense we give below, which approximates μ.

A better description of the topology on the space of measured laminations is afforded by the use of train tracks.

We first give an intuitive notion of train tracks since the formal definitions of geometric objects often obscure the rather simple pictures which they codify. To get to the punchlines quickly, I will merge the topological and metrical parts of the definition of (weighted or measured) train tracks. To obtain topological tracks, simply remove all reference to weights.

Think of a train track τ as a special type of graph embedded in a surface. The edges, here called *branches*, are assigned nonnegative weights. A vertex, called a *switch*, forces a partition on its incident branches into two classes, call them I and O. The sum of the weights on the I branches must equal that on the O branches. A closed curve in the train track, which passes through a switch, must go from an I branch to an O branch or vice versa. (This is usually stated through a requirement that the branches all be tangent at a switch and the further condition that all curves in the track be smoothly parametrized. There is no real need to look at differentiable notions.) There are two other technical conditions. One is that every simple closed curve in τ contain exactly one bivalent switch. The other is a nontriviality condition on the embedding of the track. Namely, if$R$ is a component of $S\backslash\tau$, we may double R along $\tau\backslash\{\text{switches}\}$ to obtain a punctured surface. This doubled surface is required to have negative Euler characteristic. It is fairly standard, but technical, to assume that all switches are trivalent unless they lie on a component of the track which is a simple closed curve.

A train track may be obtained from a family of disjoint, homotopically inequivalent, simple closed curves α on S by merging parallel subarcs of the curves. The isotopy class of such a family is called a *multicurve*. The weights then count the number of subarcs that have been merged to form a single branch. The switch condition means that whatever goes into a switch also must emerge from it. A branch b is called *recurrent* if there is some simple closed curve in the track which traverses b. A track is called *recurrent* if all of its branches are.

The nonrecurrent branches simply cannot carry a simple closed curve, and we will assume that such branches are not present in the tracks considered here. We may reverse the above procedure and put integral weights on the track. Then we may find a simple closed curve α on S which smashes into a multicurve in the track with the prescribed weights counting the number of times the multicurve traverses the branches. The rational-valued measures are just ratios of the number of traversings. Clear the denominators to get back the weights associated to a simple closed curve. Irrational weights are associated to limits of finite systems of disjoint simple closed curves; these are the measured geodesic laminations. Here we increase the denominators so that limits exist—more technically, we have thrown the whole numeric description into projective space.

We can use the train track to give a natural topology—in fact, a natural PL structure—to the space of systems of disjoint simple closed curves and measured geodesic laminations on a surface. A neighborhood consists of a bunch of intervals of nonnegative numbers assigned to the branches of a track. Those numbers define a range of weights for the branches. The next thing we can do with this structure is to look at *transverse recurrence*, namely, to use those weights to describe how a curve meets (transversally crosses) the track. In this way we can recover Dehn's Theorem.

I believe that train tracks and measured laminations will be enduring tools in the study of the surfaces and higher-dimensional manifolds. Thurston's compactification of the Teichmüller space T_g by measured laminations offers us a sphere at infinity to which the action of the mapping class group extends naturally and continuously. Much of our knowledge of the moduli space is conjectured by analogy with hyperbolic manifolds and proved by vastly different techniques. The hyperbolic case is often studied by lifting the given problem to the universal cover and studying the action of the cover group on the sphere at infinity (formerly called the absolute). In principle this method is now available for the moduli space using laminations. In dimension three the natural generalizations of tracks and laminations provide a description of the branched surfaces which carry incompressible surfaces (see Floyd and Oertel [5]). This provides a natural setting for the study of the object fundamental to Haken theory.

The subject matter of tracks, when suitably pruned and carefully presented, should be accessible to undergraduates. One might have to leave out some of the heavier combinatorial arguments and rely on intuition for some obvious but difficult topological results (see below).

There have been several places where measured geodesic laminations and their parametrizations using measured train tracks have been examined. The first was Thurston's semipublished manuscript [13] in which he introduced the basic ideas and interrelationships. Later, Hatcher [6] and Bleiler and Casson [2] gave nice descriptions. Penner's short introduction [11] is quite accessible. The book by Penner with Harer, which is under review here, has been circulated in draft for several years. It is the only complete treatment of the basic objects. Much of the material is a delayed writeup and expansion of Penner's dissertation. John Harer coauthored some parts of the book.

Penner's book is devoted to the static theory of tracks and their relationship to measured geodesic laminations. By static he means that the action of the mapping class groups on tracks and laminations is only outlined. Applications are left to an appendix, an epilogue, and other writings. As is typical of the monographs in the *Annals of Mathematics Studies* series, it is written at a level that is appropriate for an advanced graduate student or researchers. The difficult arguments are given in detail. The book has the main results in the static theory and will be the standard reference in the subject. It does clearly demonstrate that fairly intuitive ideas often require really tough proofs. It is possible to get an overview of the subject from a more casual reading—a quality I view as extremely desirable in any book. The book could use an index, but this flaw is noticable in many volumes in the *Annals of Mathematics Studies*.

The main emphasis in the book is on a pair of admissible moves called *splitting* and *shifting* for changing isotopy classes of tracks. The tracks that can be moved one into another form an equivalence class. In each such class there is a unique standard track. This relation was chosen so that equivalent tracks will carry the same laminations. It is then easy to give the piecewise linear structure of the space of measured geodesic laminations with compact support and, more generally, to define a natural symplectic structure on that space.

The book has an epilogue which surveys the relationship to Riemann surface theory and the action of the mapping class group. An addendum gives the main result in Penner's thesis which is a collection of explicit formulas for the action of the mapping class group in the Dehn-Thurston coordinates.

References

1. W. Abikoff, *The real analytic theory of Teichmüller space*, Lecture Notes in Math. **820** (1980), Springer-Verlag, New York.
2. A. J. Casson and S. T. Bleiler, *Automorphisms of surfaces after Nielsen and Thurston*, London Math. Soc. Student Texts, vol. **9**, Cambridge Univ. Press, Cambridge, 1988.
3. M. Dehn, *Papers on group theory and topology*, Springer-Verlag, New York, 1987.
4. A. Fathi, F. Laudenback, and V. Poenaru, *Travaux de Thurston sur les surfaces*, Astérisque **66–67** (1979).
5. W. Floyd and U. Oertel, *Incompressible surfaces via branched surfaces*, Topology **23** (1984), 117–125.
6. A. Hatcher, *Measured lamination spaces for surfaces, from the topological viewpoint*, Topology Appl. **30** (1988), 63–88.
7. L. Keen, *Intrinsic moduli on Riemann surfaces*, Ann. of Math (2) **84** (1966), 404–420.
8. G. Levitt, *Foliations and laminations on hyperbolic surfaces*, Topology **22** (1983), 119–135.
9. L. Mosher, *The classification of pseudo-anosovs*, Low Dimension Topology and Kleinian Groups, London Math. Soc. Lecture Notes, vol. 112, Cambridge Univ. Press, Cambridge, 1986, pp. 13–75.
10. S. Nag, *The complex analytic theory of Teichmüller spaces*, Wiley-Interscience, New York, 1988.
11. R. Penner, *An introduction to train tracks*, Low Dimensional Topology and Kleinian Groups, London Math. Soc. Lecture Notes, vol. 112, Cambridge Univ. Press, Cambridge, 1986, pp. 77–90.
12. C. L. Siegel, *Topics in complex function theory*, vol. 1, Wiley-Interscience, New York, 1969.
13. W. Thurston, *The geometry and topology of* 3*-manifolds*, Princeton Univ. Math. Dept., 1979.

W. Abikoff
University of Connecticut
E-mail address: abikoff@math.uconn.edu