

Contact topology and holomorphic invariants via elementary combinatorics*

Daniel V. Mathews
School of Mathematical Sciences
Monash University
Daniel.Mathews@monash.edu

Abstract

In recent times a great amount of progress has been achieved in symplectic and contact geometry, leading to the development of powerful invariants of 3-manifolds such as Heegaard Floer homology and embedded contact homology. These invariants are based on holomorphic curves and moduli spaces, but in the simplest cases, some of their structure reduces to some elementary combinatorics and algebra which may be of interest in its own right. In this note, which is essentially a light-hearted exposition of some previous work of the author, we give a brief introduction to some of the ideas of contact topology and holomorphic curves, discuss some of these elementary results, and indicate how they arise from holomorphic invariants.

Contents

1	Introduction	2
2	Symplectic and contact geometry	2
2.1	Symplectic geometry	3
2.2	Symplectic vs. complex geometry	3
2.3	Holomorphic curves in symplectic manifolds	4
2.4	Contact geometry	6
3	Quantum pawn dynamics	7
3.1	Pawns on 1-dimensional chessboards	7
3.2	Quantum “inner product”	8
3.3	Dirac sea and anti-pawns	8
3.4	The initial creation and annihilation of pawns	9
3.5	Adjoints	9
3.6	Initial creation and annihilation are adjoint	10
3.7	Adjoints and adjoints and adjoints	10
3.8	Round-up of adjoints	11
3.9	Chessboards and words	12
4	Chords on discs	12
4.1	Curves on a disc	12
4.2	Diagrams of chessboards	14
4.3	Chord diagrams of chessboards as ski slopes	15

*MSC subject classification: 57R17

4.4	Square decomposition	17
4.5	Curves on cylinders	17
4.6	Finger moves	19
4.7	“Inner product” on chord diagrams	20
4.8	Bypass discs and 120° rotations	21
4.9	A vector space of chord diagrams	23
5	Contact topology	24
5.1	Chord diagrams and contact structures	24
5.2	Contact structures and overtwisted discs	25
5.3	Contact structures on spheres and balls	26
5.4	Contact corners	27
5.5	Bypasses	28
5.6	The contact category	30
5.7	Combinatorics of contact geometry	30
6	Holomorphic invariants	31
6.1	Sutured Floer homology	31
6.2	Embedded contact homology and string homology	32
	References	33

1 Introduction

In recent years a great amount of progress has been achieved in understanding the structures of symplectic and contact geometry. Powerful invariants of manifolds and their contact and symplectic structures have been defined, such as Heegaard Floer homology and contact homology. These theories are based on generalised Cauchy-Riemann equations, holomorphic curves and their moduli spaces.

In a series of papers, the author has developed some of the structure that arises in some of the most simple cases of the holomorphic invariants known as sutured Floer homology and embedded contact homology [30, 31, 32, 33, 34]. This note is essentially a light-hearted exposition of some of that subject matter.

The structures that we discuss here are algebraic and combinatorial, and entirely elementary. They may be of interest for their own sake, as well as for their relation to other subjects, including quantum information theory, representation theory, topological quantum field theory, and especially contact geometry. That all this structure can arise from the most simple cases of sutured Floer homology and embedded contact homology, testifies to the power of these invariants.

It is our desire here to convey some of these results to a broad mathematical audience. Consequently, this note assumes no knowledge of symplectic or contact geometry or holomorphic curves. Since the subject matter is in a certain sense a “combinatorialization of contact geometry”, we will give an exposition, as we proceed, of how these elementary results relate to contact geometry — and this may serve as an unorthodox introduction to some of the ideas of contact geometry.

The reader who is only interested in combinatorics or algebra can easily skip the sections on symplectic and contact geometry and holomorphic curves. The reader without such background, but who wishes to know where the subject matter comes from, can hopefully gain here some idea of the types of considerations involved, whether in symplectic or contact geometry or holomorphic curves, and is encouraged to follow the references for further details.

2 Symplectic and contact geometry

We begin with a little — very, very little — about what symplectic and contact geometry are, and where some holomorphic invariants come from. For a proper introduction to symplectic geometry we refer to [35] and to [15] or [10] for contact geometry. We will give essentially no proofs in this section, just assert some basic facts.

None of this is needed for the combinatorics and algebra that we will shortly discuss, but it will be useful as we proceed to make connections with contact geometry and holomorphic invariants. The reader who is solely interested in combinatorial and algebraic aspects can safely skip this section.

2.1 Symplectic geometry

Symplectic geometry is the mathematical structure of Hamiltonian mechanics. A symplectic manifold is a pair

$$(M, \omega),$$

where M is a smooth manifold and ω is a closed non-degenerate differential 2-form on M . This implies that M is even-dimensional.

The key property that allows this structure to produce mechanics is that any smooth function $H : M \rightarrow \mathbb{R}$ (called a *Hamiltonian*) has a naturally associated vector field X_H on M . Namely, from H we obtain a differential 1-form dH , and then the non-degeneracy of ω gives the vector field X_H via the equation

$$\omega(X_H, \cdot) = dH.$$

(Different authors have different sign conventions in this equation, but this is the idea.)

The fact that ω is closed implies that flowing along X_H leaves ω unchanged:

$$L_{X_H}\omega = i_{X_H}d\omega + di_{X_H}\omega = i_{X_H}0 + d(dH) = 0.$$

The physical interpretation is that M is the phase space (“space of states”) of the universe, H gives the energy of any state, and X_H is the time evolution of the universe.

The simplest example of a symplectic manifold is \mathbb{R}^{2n} with the symplectic form

$$\omega = \sum_{j=1}^n dx_j \wedge dy_j.$$

Here each y_j can be considered a “position coordinate” and each x_j a corresponding “conjugate momentum”.

2.2 Symplectic vs. complex geometry

The fact that symplectic geometry only arises in even numbers of dimensions suggests a similarity to complex geometry. Indeed this is even the etymological root: Weyl introduced the word “symplectic” as a Greek version of the word “complex” [42]. The Latin *complexus* and the Greek *symplektikos* both mean “braided together”.

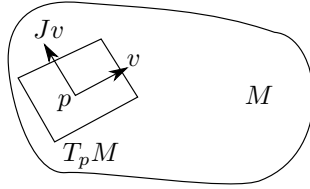
There is a difference of course. In a complex manifold every point has a neighbourhood homeomorphic to \mathbb{C}^n . In particular, for every direction there is also “ i times” that direction. This is very different from having a closed non-degenerate 2-form.

The relationship between symplectic and complex geometry has been exploited to great effect in the last 25 years or so, with the use of *holomorphic curves*, starting with the work of Gromov in 1985 [19], leading to great advances in the understanding of symplectic geometry.

In particular, any symplectic manifold (M, ω) has an *almost complex structure*. An almost complex structure is a map which is like “multiplication by i at every point of M ”. That is, J is a map

$$J : TM \rightarrow TM$$

which takes every fibre $T_p M \rightarrow T_p M$ and satisfies $J^2 = -1$.



An almost complex structure is often required to be *compatible* with the symplectic form. Roughly this means that J and ω behave like i and $\sum_j dx_j \wedge dy_j$ in \mathbb{C}^n (where the $z_j = x_j + iy_j$ are the coordinates). (Precisely, it means that $\omega(v, w) = \omega(Jv, Jw)$ for all tangent vectors v, w ; and $\omega(v, Jv) > 0$ for all tangent vectors $v \neq 0$.)

It turns out that any symplectic manifold has a compatible almost complex structure. Moreover, all compatible almost complex structures on (M, ω) are homotopic. (The space of compatible almost complex structures on (M, ω) is the space of sections of a fibre bundle over M with contractible fibres.) However, an almost complex structure by no means implies the existence of a complex structure. An almost complex structure only requires a pointwise condition. A complex structure requires local charts to \mathbb{C}^n with holomorphic transition functions, which is a far more onerous condition. To drop the “almost” requires an additional condition, the vanishing of the Nijenhuis tensor.

For proper discussion of these issues we refer to [35] or [36].

2.3 Holomorphic curves in symplectic manifolds

Gromov in [19] had the idea of considering *holomorphic curves in symplectic manifolds*. This has led to many developments, including the development of *Floer homology* theories.

The point of this note is not really to explain any such homology theories, or the detail of why they work, or how they work, or how to compute them. The point is to discuss some of the structure obtained from them in some very simple cases, which is elementary and of independent interest.

Nonetheless, for the interested reader, we can summarise some of the ideas involved, very roughly and avoiding all details, as follows. Readers not interested in holomorphic curves should skip to the next section.

- Start with a symplectic manifold (M, ω) you want to understand.
- Introduce a compatible almost complex structure J on (M, ω) . As we just noted, one always exists, and any choice is homotopic to any other.
- Take a Riemann surface (Σ, i) and consider maps

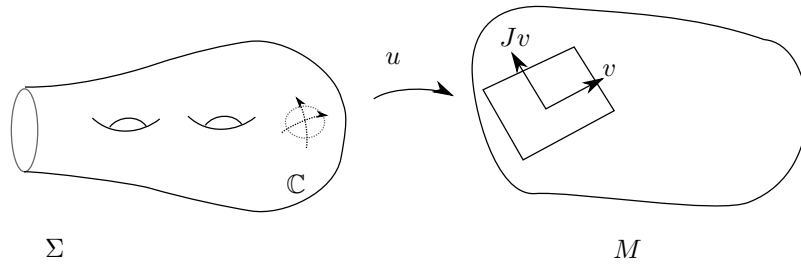
$$u : (\Sigma, i) \longrightarrow (M, J)$$

which are *holomorphic*,¹ meaning that the following diagram commutes.

$$\begin{array}{ccc} T\Sigma & \xrightarrow{Du} & TM \\ \downarrow i & & \downarrow J \\ T\Sigma & \xrightarrow{Du} & TM \end{array}$$

I.e., the *Cauchy–Riemann equations* hold: $Du \circ i = J \circ Du$.

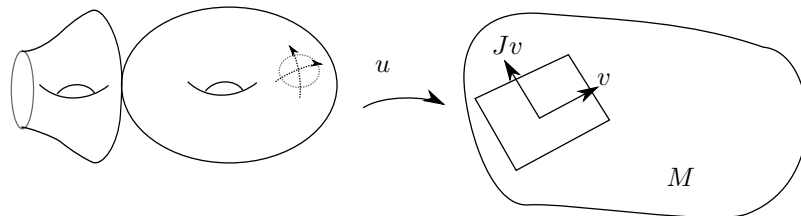
¹It is common to say “pseudo-holomorphic” or “ J -holomorphic” to indicate that the target has an almost complex structure, rather than a complex structure. We prefer simply to say holomorphic here and there should be no ambiguity.



- If we consider holomorphic curves with appropriate constraints such as marked points, fixed degree and boundary conditions, and provided sufficient transversality conditions are satisfied, then the space of holomorphic curves will be a *finite-dimensional* orbifold-like space called a *moduli space*.

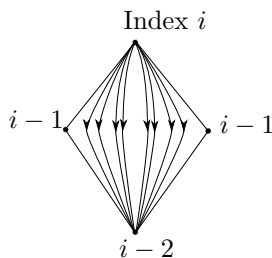
(This is a vast generalisation of elementary facts such as the following. The space of holomorphic maps $\mathbb{C}\mathbb{P}^1 \rightarrow \mathbb{C}\mathbb{P}^1$ of degree 1 fixing 3 points is a point. The space of holomorphic maps $\mathbb{C}\mathbb{P}^1 \rightarrow \mathbb{C}\mathbb{P}^1$ of degree 1 fixing 2 points is an annulus.)

- More generally, under these favourable conditions, there is an index theory for holomorphic curves. This is a version of the Riemann–Roch theorem and a special case of the Atiyah–Singer index theorem; the Cauchy–Riemann equations give a $\bar{\partial}$ operator and its index is computed in terms of topological data. The dimension of the moduli space is given in terms of the symplectic topology of the constraints. Such moduli spaces intricately encode topological data about the manifold.
- A moduli space has a natural *compactification* which is the subject of the *Gromov compactness theorem*. The compactified moduli space has a stratified boundary, and the strata are moduli spaces of “degenerate” holomorphic curves such as nodal surfaces.



- Some information can be extracted from a moduli space by exploiting the codimension-1 parts of its boundary, and defining *homology theories* based on it. Roughly, and although there are several different approaches, one takes a chain complex generated by sets of boundary conditions for holomorphic curves. One then defines a differential ∂ counting holomorphic curves between two sets of boundary conditions, at positive and negative ends respectively. Under favourable conditions, the boundary structure of the moduli space and the index theory of solutions to the Cauchy–Riemann equations give $\partial^2 = 0$.

This is analogous to how the singular homology of a smooth manifold can be obtained from a Morse function f via a chain complex (the *Morse complex*) generated by critical points of f , and a differential counting gradient trajectories between critical points. (Critical points are “boundary conditions for gradient trajectories”).



- Depending on the details of how this chain complex is constructed, one can obtain various flavours of *Floer homology*, *contact homology* or *symplectic field theory*. An intricate algebraic structure of generating functions can be used to keep track of detail about counts of holomorphic curves.
- It often turns out that the resulting homology is independent of the choice of almost complex structure J , and sometimes even of the underlying symplectic or contact structure. It is possible to obtain *smooth* manifold or *knot* invariants.
- *Heegaard Floer homology theories* are a very powerful variant of the above ideas, introduced by Ozsváth and Szabó [39, 40, 41]. From a 3-manifold M , we take its *Heegaard decomposition* consisting of a surface Σ with curves $\alpha_1, \dots, \alpha_g$ bounding discs on one side and β_1, \dots, β_g on the other. We then (as in [27]) consider holomorphic curves in the symplectic manifold $\Sigma \times I \times \mathbb{R}$ with asymptotics prescribed by the α_i and β_j .

It turns out that one can set up a chain complex so that the resulting homology is independent of any symplectic or almost complex structure chosen, giving a *smooth manifold invariant* of M . Different versions, such as the *hat*, *infinity*, *minus*, *knot*, *bordered* [28] and *sutured* [26] Heegaard Floer homologies vary between the types of manifolds they apply to, the complexity of the theory, and the ease of obtaining information.

One indication of the power of Heegaard Floer homology is that it detects the genus of a knot [38, 37]; this can be computed combinatorially [29].

We shall say no more about holomorphic curves until the final section, when we briefly return to sutured Floer homology to see how it is related to the elementary structures forming the main subject of this note.

2.4 Contact geometry

Contact geometry is the odd-dimensional sibling of symplectic geometry. Its roots can be traced back to Lie’s work on systems of differential equations, and arguably back to Christiaan Huygens [13, 14].

One way to define a symplectic structure is a 2-form ω satisfying

$$d\omega = 0 \quad \text{and} \quad \omega^n \text{ is nowhere zero, i.e. a volume form.}$$

Analogously, a *contact form* α on a $(2n + 1)$ -dimensional manifold M is a 1-form satisfying

$$\alpha \wedge (d\alpha)^n \text{ is nowhere zero, i.e. a volume form.}$$

The kernel of α is a hyperplane field ξ on M and it is this plane field that is called a *contact structure*.

Contact manifolds arise in considering wavefronts in optics. They also arise naturally as submanifolds of symplectic manifolds. In fact, so many geometric problems can be interpreted in terms of contact geometry, that Arnold once famously said that “contact geometry is all geometry”. [1]

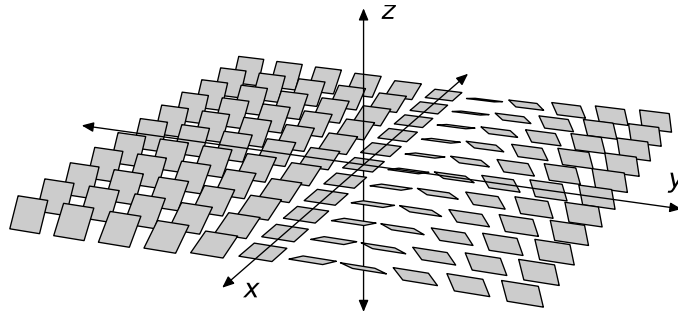
The condition $\alpha \wedge (d\alpha)^n \neq 0$ has a geometric interpretation from Frobenius’ theorem in differential geometry. It is that ξ is *totally non-integrable*.

We are interested in 3-dimensional contact manifolds. Non-integrability then means that there is no 2-dimensional surface immersed in M which is tangent to ξ . There are 1-dimensional curves tangent to ξ , but not 2-dimensional surfaces, not even locally.

The simplest example of a 3-dimensional contact manifold is \mathbb{R}^3 with contact form

$$\alpha = dz - ydx.$$

The corresponding contact structure $\xi = \ker \alpha$ is shown below.



Any contact manifold with a contact form (M, α) can be used to obtain a symplectic manifold, called its *symplectization*:

$$(M \times \mathbb{R}, d(e^t \alpha)),$$

where t is the coordinate on \mathbb{R} . It is then possible to consider holomorphic curves in this symplectization, with asymptotics prescribed by contact geometry. Using the analytic techniques mentioned above, we can obtain homology theories such as *contact homology* [2] and *embedded contact homology* [25].

Having said a little about contact and symplectic topology, I now propose to drop the subject entirely and talk about some seemingly unrelated algebra and combinatorics. We will later see how this becomes, in a certain sense, a combinatorial version of contact topology, and how it is related to holomorphic invariants.

3 Quantum pawn dynamics

The theory of quantum pawn dynamics, or QPD, is a strange theory of pawns on chessboards.

3.1 Pawns on 1-dimensional chessboards

This theory has no space, no time, and no proper chessboards. It does have pawns though. The pawns move along a finite 1-dimensional chessboard.



The pawns move from left to right. As they are pawns, they can only move ahead one space at a time, and only into an unoccupied square. There is no capturing, no en passant, and no double first moves. Two pawns cannot occupy the same square. So from the situation above, the pawns could eventually arrive at



in two moves, but could never get to



as the middle pawn would have to move backwards.

There is nothing special about the number of pawns or the size of the chess-“board”. We can have a chess-“board” of any length n , and we can have any number of pawns n_p where $0 \leq n_p \leq n$. Any setup of the board is a state of the QPD universe.

3.2 Quantum “inner product”

We now have pawns, and they move — hence, dynamics. We now make the dynamics quantum by declaring an “inner product” $\langle \cdot | \cdot \rangle$. In quantum mechanics the inner product is supposed to give you a probability amplitude for getting from one state of the universe to another.

The QPD universe is more binary than that. Given one setup of pawns w_0 , you can either get to another setup of pawns w_1 or you can’t. And in fact, we do not need a number system more complicated than \mathbb{Z}_2 . So we do not have probabilities so much as *possibilities*.

We declare that

$$\langle w_0 | w_1 \rangle = \begin{cases} 1 & \text{if it is possible for pawns to move from } w_0 \text{ to } w_1 \\ & \text{(in some number of moves, possibly 0);} \\ 0 & \text{if not.} \end{cases}$$

So our examples above translate to

$$\left\langle \begin{array}{|c|c|c|c|c|c|} \hline \text{♟} & & \text{♟} & \text{♟} & & \\ \hline \end{array} \mid \begin{array}{|c|c|c|c|c|c|} \hline & & \text{♟} & \text{♟} & & \text{♟} \\ \hline \end{array} \right\rangle = 1$$

and

$$\left\langle \begin{array}{|c|c|c|c|c|c|} \hline \text{♟} & & \text{♟} & \text{♟} & & \\ \hline \end{array} \mid \begin{array}{|c|c|c|c|c|c|} \hline \text{♟} & \text{♟} & & & \text{♟} & \\ \hline \end{array} \right\rangle = 0$$

Moreover, this is *quantum* pawn dynamics, so we can have *superpositions* of states — entangled chessboards! We consider that we can *add* chessboards. But, as we said, we will not consider numbers more complicated than 1; we take coefficients in \mathbb{Z}_2 . The space of states is the \mathbb{Z}_2 -vector space freely generated by chessboard setups. We declare that the “inner product” is bilinear, e.g.:

$$\left\langle \begin{array}{|c|c|c|c|c|c|} \hline \text{♟} & & \text{♟} & \text{♟} & & \\ \hline \end{array} \mid \begin{array}{|c|c|c|c|c|c|} \hline & & \text{♟} & \text{♟} & & \text{♟} \\ \hline \end{array} + \begin{array}{|c|c|c|c|c|c|} \hline \text{♟} & \text{♟} & & & \text{♟} & \\ \hline \end{array} \right\rangle = 1+0 = 1.$$

Importantly, note that the “inner product” is *not symmetric*. But it is asymmetric in a very interesting way, as we’ll see. It is a “booleanized” *partial order* on the setups of the chessboard. In fact, the configurations of a chessboard form a *complete lattice* in a natural way.

3.3 Dirac sea and anti-pawns

Actually, there is some strong symmetry in QPD. We can think of chessboard with no pawns as a thriving sea of anti-pawns. We can think of any square not occupied by a pawn, as containing an anti-pawn. An anti-pawn is just an “absence of pawn”. This is very similar to the idea of the “Dirac sea” of matter and anti-matter. We draw pawns as white and anti-pawns as black.

$$\begin{array}{|c|c|c|c|c|c|} \hline \text{♟} & & \text{♟} & \text{♟} & & \\ \hline \end{array} = \begin{array}{|c|c|c|c|c|c|} \hline \text{♟} & \text{♚} & \text{♟} & \text{♟} & \text{♚} & \text{♚} \\ \hline \end{array}$$

Note when a pawn moves right, an anti-pawn moves left. So we can get from one setup of the chessboard to another, iff all pawns move right, iff all anti-pawns move left.

So white and black are not exactly like the opposing sides of chess, but they do move in opposite directions. Each is the absence of the other.

In each of the examples above, we imagine we have 6 pawn-particles: 3 pawns, and 3 anti-pawns. With n as the number of squares on the chessboard / number of pawn-particles, and n_p the number of pawns, we let $n_q = n - n_p$ be the number of anti-pawns.

3.4 The initial creation and annihilation of pawns

Being a quantum field theory of pawns, QPD will have to allow us to create and annihilate pawns and anti-pawns.

We'll define the *initial pawn creation operator* $a_{p,0}^*$ to take a chessboard setup, and adjoin a new *initial* square to the chessboard — i.e. at the left hand side. This new square/particle has a pawn on it (is a pawn-particle).

$$a_{p,0}^* \begin{array}{|c|c|c|c|c|c|} \hline \text{♞} & \text{♜} & \text{♞} & \text{♞} & \text{♜} & \text{♜} \\ \hline \end{array} = \begin{array}{|c|c|c|c|c|c|} \hline \text{♞} & \text{♞} & \text{♜} & \text{♞} & \text{♞} & \text{♜} & \text{♜} \\ \hline \end{array}$$

We'll also define the *initial pawn annihilation operator* $a_{p,0}$ to delete an initial pawn. That is, it will annihilate the leftmost square from the chessboard — the chessboard shrinks and a particle disappears — provided that square contains a pawn. What does the QPD universe do if the initial square does not contain a pawn, but an anti-pawn? Why, it returns an error of course: error 404 universe not found mod 2 is 0.

So, for example:

$$a_{p,0} \begin{array}{|c|c|c|c|c|c|} \hline \text{♞} & \text{♜} & \text{♞} & \text{♞} & \text{♜} & \text{♜} \\ \hline \end{array} = \begin{array}{|c|c|c|c|c|} \hline \text{♜} & \text{♞} & \text{♞} & \text{♜} & \text{♜} \\ \hline \end{array}$$

$$a_{p,0} \begin{array}{|c|c|c|c|c|} \hline \text{♜} & \text{♞} & \text{♞} & \text{♜} & \text{♜} \\ \hline \end{array} = 0.$$

The vacuum \emptyset is the state of the QPD universe with no chessboard. (This is different from 0.) The universe is a lonely, empty place without a chessboard. But you can make a bang and start your universal chessboard by creating an initial pawn.

$$a_{p,0}^* \emptyset = \begin{array}{|c|} \hline \text{♞} \\ \hline \end{array}$$

Everything we have said for pawns applies also to anti-pawns. So there is an *initial anti-pawn creation operator* $a_{q,0}^\dagger$, which creates a new leftmost square with an anti-pawn. There is also an *initial anti-pawn annihilation operator* $a_{q,0}$, which annihilates a leftmost anti-pawn, or else returns error 0.

Using initial pawn and anti-pawn creation operators you can build any chessboard setup out of nothing.

$$a_{p,0}^* a_{q,0}^\dagger a_{p,0}^* a_{q,0}^\dagger a_{p,0}^* a_{q,0}^\dagger a_{q,0}^\dagger \emptyset = \begin{array}{|c|c|c|c|c|c|} \hline \text{♞} & \text{♜} & \text{♞} & \text{♞} & \text{♜} & \text{♜} \\ \hline \end{array}$$

The $*$ and \dagger on creation operators refer to *adjoints*, which we discuss next. In the world of lattices and partial orders these are known as *Galois connections*. (See [31] for further details.)

3.5 Adjoints

It's standard notation in physics to write annihilation operators with an a and creation operators as a^* or a^\dagger . And these operators are supposed to be related: they are *adjoint* with respect to the inner product.

$$\langle ax|y \rangle = \langle x|a^*y \rangle, \quad \langle x|ay \rangle = \langle a^*x|y \rangle.$$

$$\begin{aligned}
&= \left\langle \begin{array}{|c|c|c|c|c|c|} \hline \text{♙} & & \text{♘} & \text{♘} & & \\ \hline \end{array} \mid \begin{array}{|c|c|c|c|c|c|} \hline & \text{♙} & \text{♘} & & \text{♘} & \\ \hline \end{array} \right\rangle \\
&= \left\langle \begin{array}{|c|c|c|c|c|c|} \hline & \text{♘} & \text{♘} & & & \\ \hline \end{array} \mid \begin{array}{|c|c|c|c|c|c|} \hline & \text{♘} & & \text{♘} & & \\ \hline \end{array} \right\rangle = 1
\end{aligned}$$

We conclude that $a_{p,0}^{**}$ is the operator which *annihilates the first pawn on the chessboard*, from left to right. We'll write this as $a_{p,1}$.

But why stop there? What is $a_{p,1}^*$? This is the operator f which satisfies

$$\langle a_{p,1}x|y \rangle = \langle x|fy \rangle.$$

Given chessboards x and y , you delete the first pawn from x , and compare $a_{p,1}x$ and y . You now want to come up with a chessboard fy , so that comparing the deleted chessboard $a_{p,1}x$ to y always gives the same result as comparing x to fy . Well, you certainly don't want to disturb the relative positions of the pawns. You just want to quietly slip a first pawn into y in such a way that the first pawn of x can easily move there. And the place to slip in that pawn is to *double the first pawn in y* .

$$\begin{aligned}
&\left\langle a_{p,1} \begin{array}{|c|c|c|c|c|c|} \hline & \text{♘} & & \text{♘} & & \text{♘} \\ \hline \end{array} \mid \begin{array}{|c|c|c|c|c|c|} \hline & & & \text{♘} & & \text{♘} \\ \hline \end{array} \right\rangle \\
&= \left\langle \begin{array}{|c|c|c|c|c|c|} \hline & & \text{♘} & & & \text{♘} \\ \hline \end{array} \mid \begin{array}{|c|c|c|c|c|c|} \hline & & & \text{♘} & & \text{♘} \\ \hline \end{array} \right\rangle \\
&= \left\langle \begin{array}{|c|c|c|c|c|c|} \hline & \text{♙} & & \text{♘} & & \text{♘} \\ \hline \end{array} \mid \begin{array}{|c|c|c|c|c|c|} \hline & & & \text{♙} & \text{♘} & \text{♘} \\ \hline \end{array} \right\rangle
\end{aligned}$$

3.8 Round-up of adjoints

We can continue in this fashion, computing more and more adjoints. Given what's above, it might not be too surprising to discover that the iterated adjoints of $a_{p,0}$ can be expressed as

$$a_{p,0} \rightarrow a_{p,0}^* \rightarrow a_{p,1} \rightarrow a_{p,1}^* \rightarrow a_{p,2} \rightarrow \cdots \rightarrow a_{p,n_p} \rightarrow a_{p,n_p}^* \rightarrow a_{p,\Omega} \rightarrow a_{p,\Omega}^*$$

where the arrows represented the operation of taking the $*$ -adjoint. The operators $a_{p,i}$ for $1 \leq i \leq n_p$ delete the i 'th pawn. The operators $a_{p,i}^*$ double the i 'th pawn. And at the end we obtain *final* creation and annihilation operators, which we've denoted $a_{p,\Omega}^*$ and $a_{p,\Omega}$. These are to the right end of a chessboard what initial creation and annihilation operators are to the left end.

We can do just the same for the anti-pawn creation and annihilation operators. Drawing a similar diagram, with arrows representing the $*$ adjoint (the inverse of the \dagger adjoint), we obtain something similar (although in a different direction).

$$a_{q,\Omega}^\dagger \rightarrow a_{q,\Omega} \rightarrow a_{q,n_q}^\dagger \rightarrow a_{q,n_q} \rightarrow \cdots \rightarrow a_{q,2} \rightarrow a_{q,1}^\dagger \rightarrow a_{q,1} \rightarrow a_{q,0}^\dagger \rightarrow a_{q,0}$$

It follows that

$$a_{p,0}^{*2n_p+2} = a_{p,\Omega} \quad \text{and} \quad a_{q,0}^{\dagger 2n_q+2} = a_{q,\Omega}.$$

But there's no reason to stop there; you can just keep taking adjoints. The description of these adjoints in terms of pawns, however, becomes more complicated.

It turns out that these adjoints are *periodic*.

Theorem 3.1. *On a chessboard with n squares,*

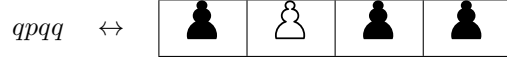
$$a_{p,0}^{*2n+2} = a_{p,0}, \quad a_{q,0}^{\dagger 2n+2} = a_{q,0}.$$

We might say that " $*^{2n+2} = 1$ ". Our proof of this involves a little more geometry, by considering the combinatorics of chords on discs, as discussed in the next section.

The various creation and annihilation operators in different positions actually satisfy the relations of a *simplicial set*: see [31] for details.

3.9 Chessboards and words

By now “chessboard notation” is becoming somewhat unwieldy. In fact, we can denote any chessboard by a sequence of p 's and q 's, where a p is a pawn and q an anti-pawn.



$$\langle pqpppq | qpqpq \rangle = \left\langle \begin{array}{|c|c|c|c|c|c|} \hline \text{white pawn} & & \text{white pawn} & \text{white pawn} & & \\ \hline \end{array} \mid \begin{array}{|c|c|c|c|c|c|} \hline & \text{white pawn} & \text{white pawn} & & \text{white pawn} & \\ \hline \end{array} \right\rangle = 1.$$

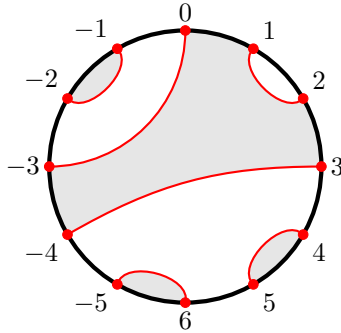
4 Chords on discs

4.1 Curves on a disc

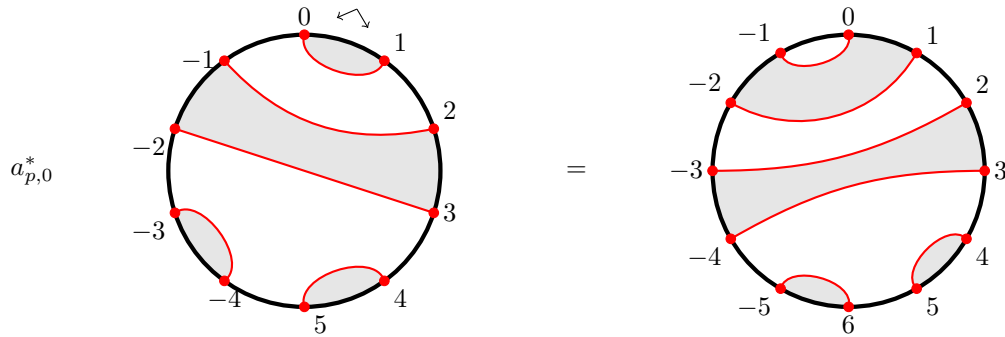
We now consider curves on discs and cylinders.

We consider a disc with $2n + 2$ points marked on the boundary, which we number clockwise by integers modulo $2n + 2$. The point numbered 0 is considered a basepoint. We consider *chord diagrams*, which are collections of non-intersecting curves joining the points up in pairs. We only consider them up to isotopy relative to the boundary.

Note that each chord in a chord diagram must connect points with opposite parity. We can shade the complementary regions of a chord diagram, so that each boundary interval $(2i, 2i + 1)$ is shaded black, and each interval $(2i - 1, 2i)$ is shaded white.

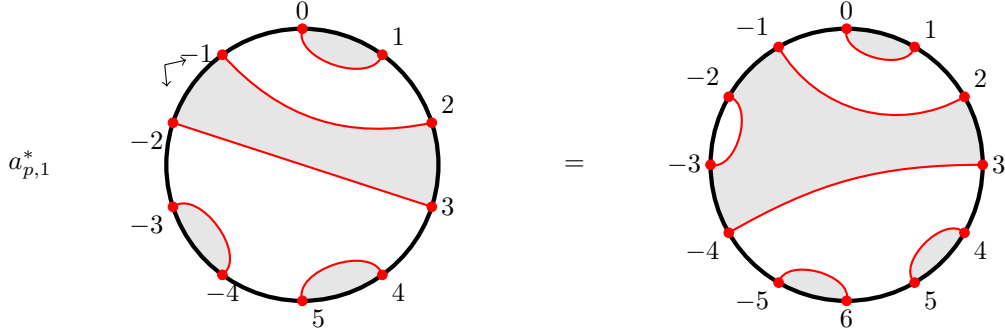


Now we will declare that the effect of the *creation operator* $a_{p,0}^*$ on a chord diagram is to insert a new chord between the points currently numbered 0 and 1; the new points are numbered -1 and 0 and so the points on the left need to be renumbered. The existing chords remain in place but are pushed around the disc as shown.



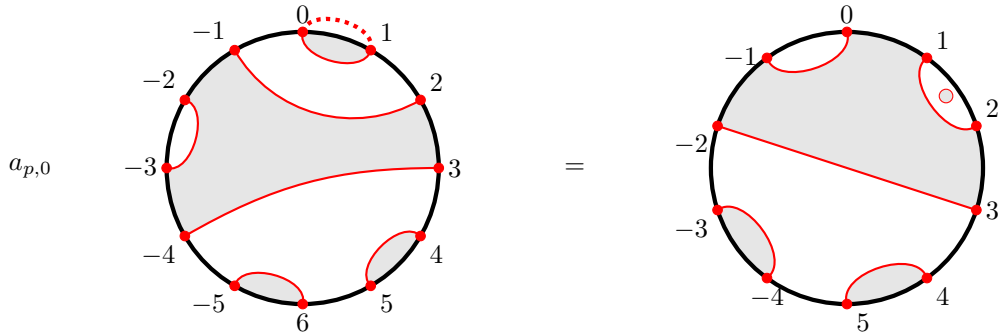
More generally, we will declare that the creation operator $a_{p,i}^*$ inserts a new chord between the points currently labelled $-2i$ and $-2i + 1$; the new points are labelled $-2i - 1$ and $-2i$, and points are renumbered accordingly,

For example



We will declare the effect of the *annihilation operator* $a_{p,0}$ on a chord diagram to “close off” the points 0 and 1 by joining them with a chord, and pushing this into the disc. Points numbered 3, 4, ... retain their numbering; the points labelled $-1, -2, \dots$ have their label increased by 2. We obtain a diagram with two fewer points on the boundary. It might have a closed curve as well as chords.

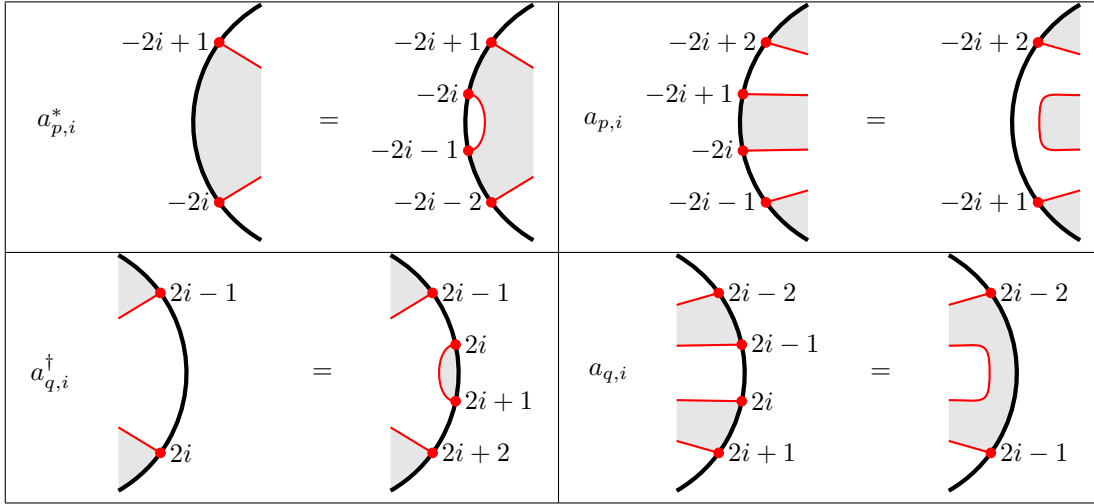
For example,



More generally, we declare that the annihilation operator $a_{p,i}$ closes off the points $-2i$ and $-2i + 1$, and relabels points: the points $-2i + 2, -2i + 3, \dots$ retain their numbering but the points $-2i - 1, -2i - 2, \dots$ have their label increased by 2. Again we might see a closed curve as a result.

When $i = 0$, both $a_{p,i}^*$ and $a_{p,i}$ perform operations near the basepoint labelled 0. When i increases by 1, those operations occur 2 spots anticlockwise.

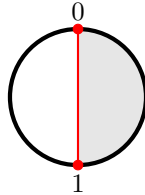
We will also declare creation operators $a_{q,i}^\dagger$ and annihilation operators $a_{q,i}$. These have a similar effect but on the other side of the diagram. We declare $a_{q,i}^\dagger$ inserts a new chord between the points labelled $2i - 1$ and $2i$; and we declare that $a_{q,i}$ closes off the points $2i - 1$ and $2i$. So $a_{q,0}$ and $a_{q,0}^\dagger$ perform operations near the basepoint, and when i increases by 1, those operations occur 2 spots clockwise.



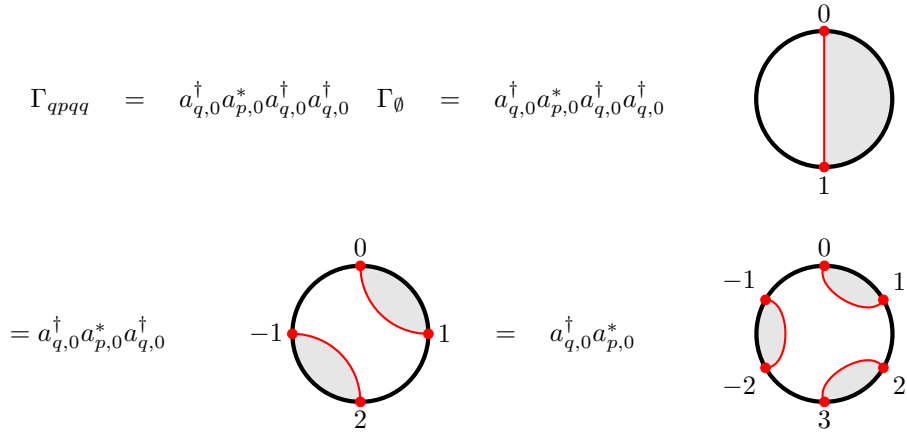
Note that the p -creation operator always creates a new white region, and the q -creation operator always creates a new black region. Also, the p -annihilation operator always closes off a black region, while the q -annihilation operator always closes off a white region. The similarity to pawn colours is not coincidental.

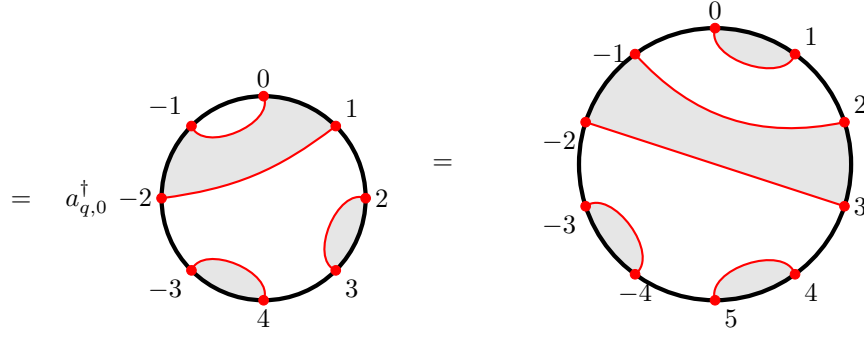
4.2 Diagrams of chessboards

We'll call the simplest possible chord diagram, with one chord, *the vacuum* Γ_\emptyset .

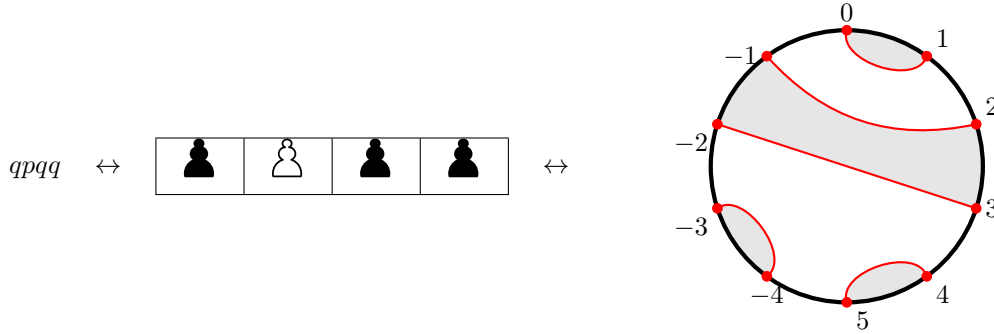


Now for any word w in the letters p and q , we can apply a corresponding sequence of initial creation operators $a_{p,0}^*$ and $a_{q,0}^\dagger$ to the vacuum chord diagram Γ_\emptyset to get a chord diagram Γ_w for the word w . For example, for the word $w = pqqq$ we obtain





Now a word w in p and q corresponds to a chessboard, where each p stands for a pawn and each q stands for an anti-pawn. So we can say that the chessboard w has chord diagram Γ_w .



It turns out that the creation and annihilation operators $a_{p,i}, a_{p,i}^*, a_{q,i}, a_{q,i}^\dagger$ act on chessboards and chord diagrams coherently.

Proposition 4.1. For any chessboard/word w ,

$$\Gamma_{a_{p,i}^* w} = a_{p,i}^* \Gamma_w.$$

In other words, the following diagram commutes.

$$\begin{array}{ccc}
 & \text{Draw} & \\
 & \text{diagram} & \\
 w & \rightsquigarrow & \Gamma_w \\
 \text{Create} \downarrow & & \downarrow \text{Create} \\
 \text{pawn} & \text{Draw} & \text{chord} \\
 & \text{diagram} & \\
 a_{p,i}^* w & \rightsquigarrow & \Gamma_{a_{p,i}^* w}
 \end{array}$$

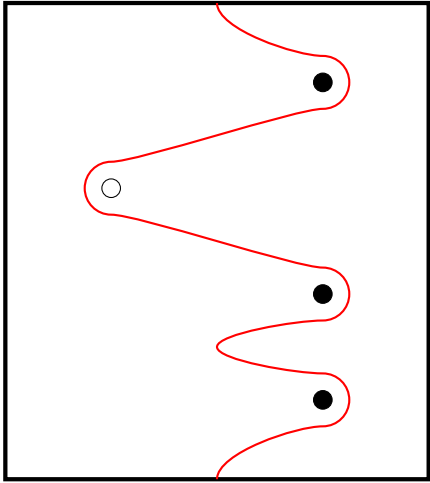
A similar result holds for operators $a_{q,i}^\dagger$. And also for annihilation operators $a_{p,i}, a_{q,i}$, provided we take any diagram with a closed loop to be zero. Annihilating a pawn in the wrong place gives an error “universe not found”; annihilating a chord in the wrong place, so that the result is not a chord diagram, gives an error “chord diagram not found”.

4.3 Chord diagrams of chessboards as ski slopes

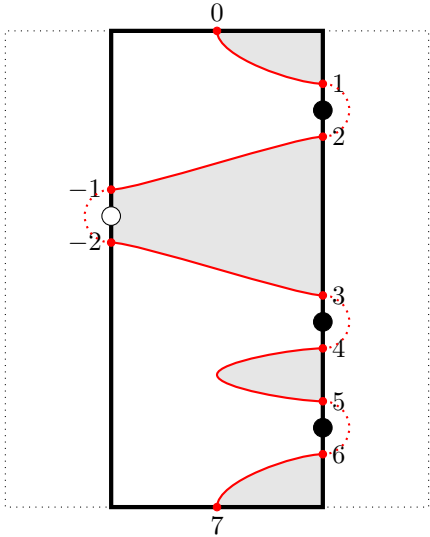
In fact there’s a quicker way to draw the diagram of a chessboard, which also illustrates why the above proposition is true. But we need to consider a different sport: skiing.

We imagine we have a ski slope, where as usual we ski from top to bottom. It is rectangular in shape, with a starting line at the top, a finishing line at the bottom, and a left and right side along the slope.

We consider a slalom run. There are obstacles, which we can imagine as poles in the ground. As you go down the slope, you have to round all the obstacles, from top to bottom; and after rounding each one you have to return to the centre of the course. There are some (white) poles placed on the left, and some (black) on the right of the slope. We'll write p for a pole on the left, and q for a pole on the right. For the slalom course $qpqq$ the course looks as follows.



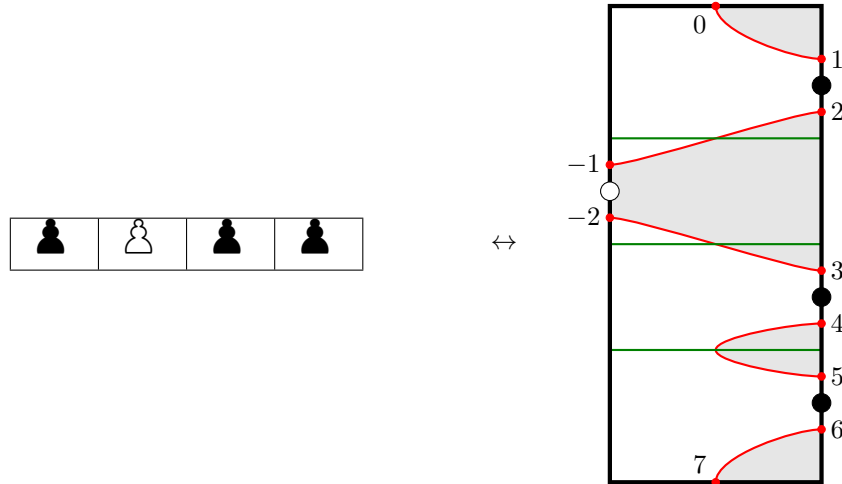
Now, if we just consider that part of the course inside the obstacles, we get an interesting chord diagram. In fact, it is just the chord diagram Γ_{qpqq} .



It's not difficult to see why this skiing algorithm gives the correct chord diagram for each word/chessboard. Having the chord diagram arranged this way, with all the chords coming from pawns / p 's on the left, and all the chords coming from anti-pawns / q 's on the right, makes it easier to see why the chord diagram operations correspond to operations on chessboards. For instance, $a_{p,i}^*$ adds an extra component to the ski run, entering from the left side, then sharply turning back to the left hand side of the slope. This is precisely what you get when you add an extra pole on the left.

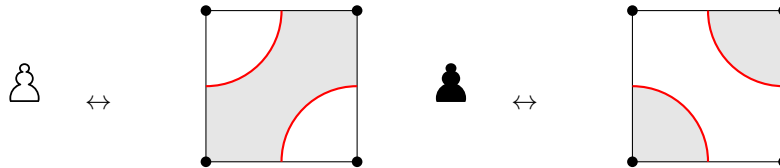
4.4 Square decomposition

We can also note that the pawns and anti-pawns correspond to a precise decomposition of the chord diagram into squares, as shown by the green lines.



One can essentially read the chord diagram as a chessboard in this way: reading the ski slope from top to bottom, reads off the chessboard from left to right.

Moreover, we see that each of these squares in the chord diagram contains chords in a particular configuration, corresponding to a pawn or anti-pawn.

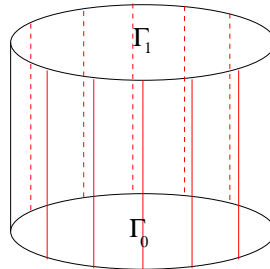


In this way, each pawn or anti-pawn provides one of two ways of drawing in chords into the chord diagram — providing, in a sense, *one bit of information* and suggesting relations to quantum information theory. Moreover, since we have creation and annihilation operators for pawns and anti-pawns, and they can also be regarded as bits of information, one is reminded of the “it from bit” idea of John Archibald Wheeler [43]. These ideas are more fully developed in [33].

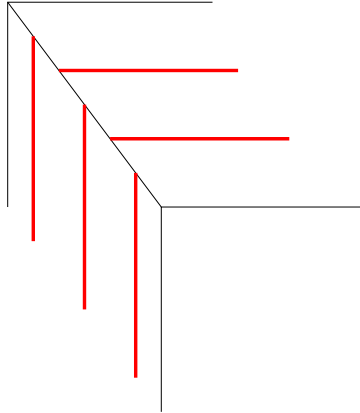
This decomposition of a chord diagram into pieces, each of which has curves in one of two specified configurations, is also reminiscent of statistical mechanics.

4.5 Curves on cylinders

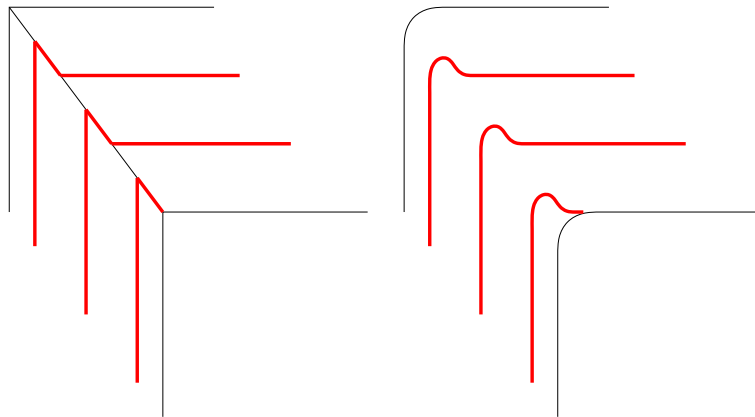
Consider the cylinder shown. Its boundary consists of discs on the top and bottom, and a vertical annulus. On the vertical annulus we have some vertical curves, drawn in red.



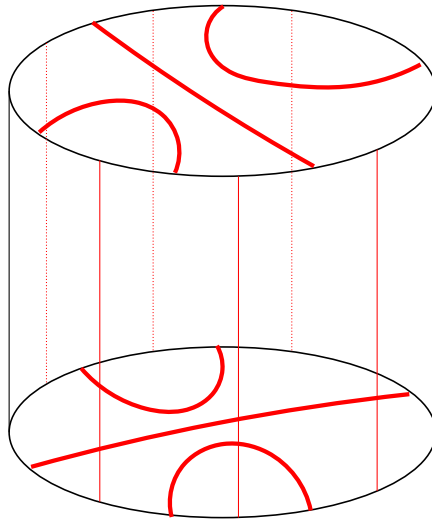
We are going to draw chord diagrams Γ_0 and Γ_1 on the bottom and top discs, and then join up all the curves to obtain curves on the boundary of the cylinder, which of course is topologically a sphere. However, when we do so, we draw the curves are arranged along the corners as shown:



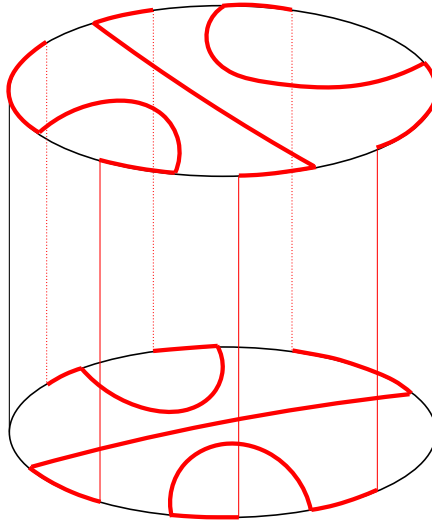
When we connect up the curves, we do it in the following fashion. We could imagine that we are turning and walking along the corner, as shown on the left; or rounding the corners and the curves, as shown on the right.



Here we have an example, with a chord diagram at the top (namely Γ_{qp} , if we regard the basepoint as being at the back), and another chord diagram at the bottom (namely Γ_{pq}).

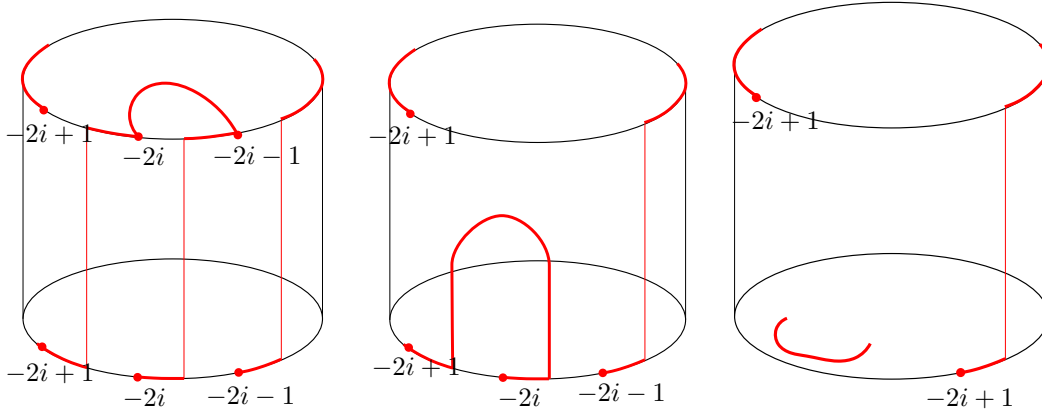


If we join up the curves as we specified, we obtain the following, which you can check is a single connected curve on the cylinder/sphere. (A fun exercise is to prove that if we draw the *same* chord diagram on the top and bottom of the cylinder, aligned exactly, then joining up the curves on the cylinder in this fashion always gives a single connected curve.)

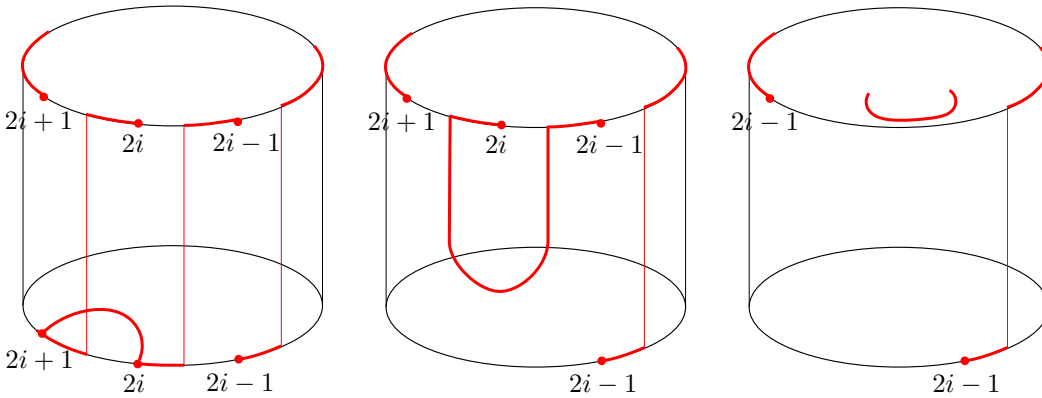


4.6 Finger moves

Now, suppose we perform an annihilation operator $a_{p,i}$ on the chord diagram on the bottom of the cylinder. The topological effect is the same as performing the creation operator $a_{p,i}^*$ on the top of the cylinder — the two are related by a *finger move* as shown.



Similarly, if we perform the creation operator $a_{q,i}^\dagger$ on the *bottom* of the cylinder, the topological effect is the same as performing the annihilation operator $a_{q,i}$ on the *top*.



So these creation and annihilation operators, inserting and closing off chords in a chord diagram, are related in a way that is quite similar to an adjoint.

4.7 “Inner product” on chord diagrams

In fact, based on the above, we will define an “inner product” of two chord diagrams $\langle \Gamma_0 | \Gamma_1 \rangle$ via what happens when you insert those two diagrams into the cylinder and round the corners. The result must be a collection of curves on the sphere; it may be one connected curve, or it may be disconnected and consist of several connected curve components.

We set

$$\langle \Gamma_0 | \Gamma_1 \rangle = \begin{cases} 1 & \text{if the result of rounding the curves on the cylinder} \\ & \text{is a single connected curve on the sphere;} \\ 0 & \text{if the result is disconnected.} \end{cases}$$

Note the asymmetry: we require that Γ_0 go on the bottom, and Γ_1 on the top.

The example drawn above in section 4.5 shows that

$$\langle \Gamma_{pq} | \Gamma_{qp} \rangle = 1.$$

The finger move we drew above shows that

$$\langle a_{q,i}^\dagger \Gamma_0 | \Gamma_1 \rangle = \langle \Gamma_0 | a_{q,i} \Gamma_1 \rangle.$$

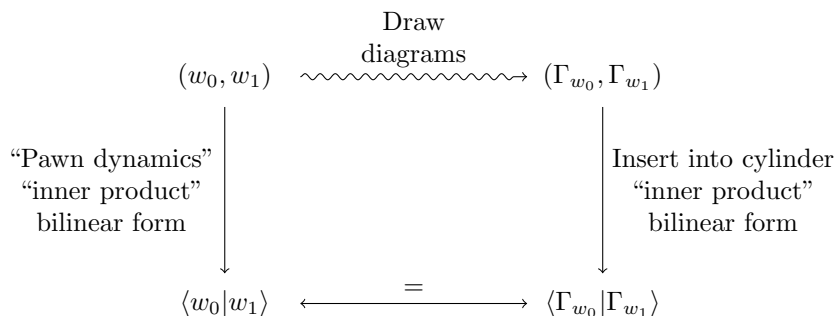
Similar finger moves show that annihilation and creation operators on chord diagrams are adjoint — just as on chessboards.

$$\begin{aligned}\langle a_{p,i}\Gamma_0|\Gamma_1\rangle &= \langle \Gamma_0|a_{p,i}^*\Gamma_1\rangle \\ \langle a_{q,i+1}\Gamma_0|\Gamma_1\rangle &= \langle \Gamma_0|a_{q,i}^\dagger\Gamma_1\rangle \\ \langle a_{p,i}^*\Gamma_0|\Gamma_1\rangle &= \langle \Gamma_0|a_{p,i+1}\Gamma_1\rangle.\end{aligned}$$

One can prove that this bilinear form is isomorphic to the one defined on chessboards.

Theorem 4.2 ([30, 31]). *For any two chessboards/words w_0, w_1 ,*

$$\langle w_0|w_1\rangle = \langle \Gamma_{w_0}|\Gamma_{w_1}\rangle.$$

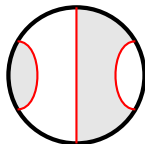


That is, the property of pawns on a chessboard being able to move from one setup to another, is precisely the property of curves on the cylinder joining up to give a single connected curve.

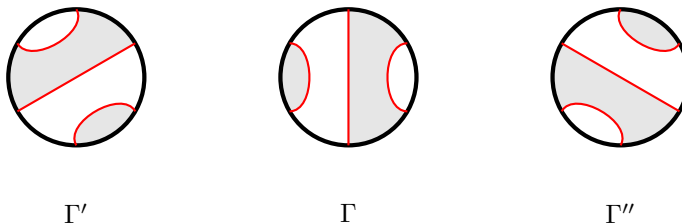
Having seen this, it is perhaps more plausible why repeated adjoints of chessboard operators might be periodic.

4.8 Bypass discs and 120° rotations

Suppose we have a chord diagram Γ on a disc D , and we consider a sub-disc $B \subset D$ on which the chord diagram appears as shown.

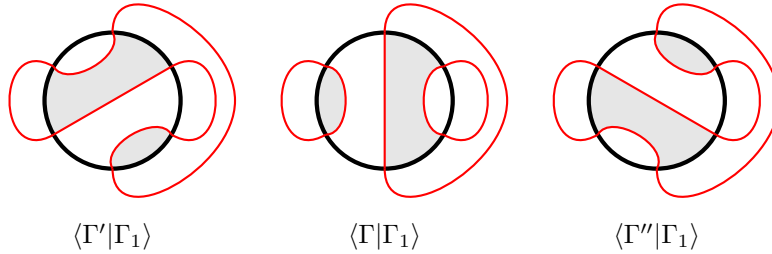


Such a disc, on which the chord diagram restricts to 3 parallel arcs, is called a *bypass disc*. There are two natural ways to adjust this chord diagram, consistent with the colours, which amount to 120° rotations. These operations are called *bypass surgeries*.



With these surgeries on $B \subset D$, we have three chord diagrams on D , which we denote $\Gamma, \Gamma', \Gamma''$. We consider inserting $\Gamma, \Gamma', \Gamma''$ successively into one end of a cylinder, with some other chord diagram Γ_1 on the other end. We obtain three sets of curves on the cylinder.

Below is drawn a possible arrangement of curves on the cylinder. We draw what happens inside B inside a circle, and what happens outside B outside the circle — this does not necessarily correspond to the top and bottom of the cylinder.



We find here that, of the three sets of curves on the cylinder, 2 of them are connected. The key observation is that *whatever chord diagrams we have for Γ and Γ_1 , out of the 3 sets of curves obtained on the cylinder, the number of sets of curves which is connected is even.*

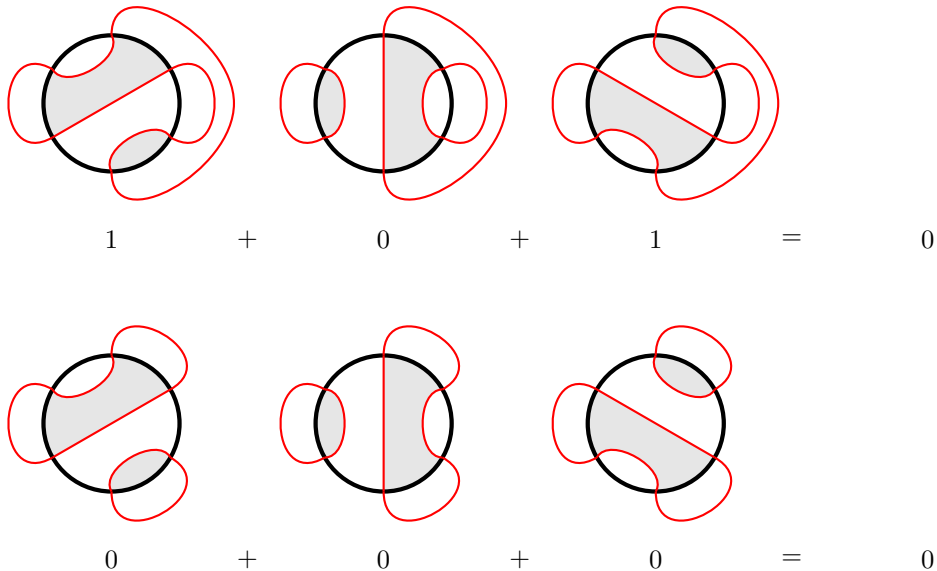
Proposition 4.3. *Let Γ, Γ_1 be chord diagrams on the disc D with the same number of chords. Then, with Γ', Γ'' obtained from Γ as above,*

$$\langle \Gamma | \Gamma_1 \rangle + \langle \Gamma' | \Gamma_1 \rangle + \langle \Gamma'' | \Gamma_1 \rangle = 0.$$

In other words,

$$\langle \Gamma + \Gamma' + \Gamma'' \mid \Gamma_1 \rangle = 0.$$

Proof. The arrangement is either as in the above diagram, giving $1 + 0 + 1 = 0 \pmod{2}$, or more degenerate, giving $0 + 0 + 0 = 0$. See below.



□

4.9 A vector space of chord diagrams

We have seen how some chord diagrams can be drawn from chessboards. But this is a very small class of chord diagrams; there are many more chord diagrams than chessboard / ski slope chord diagrams.

On the other hand, our creation and annihilation operators are defined on any chord diagram, not just chessboard ones. It's easy to draw in extra chords, or close off chords, on any chord diagram, not just those drawn from chessboards. The adjoint relations from finger moves also work generally, and the bilinear form / "inner product" defined from insertion into a cylinder also works for any chord diagram.

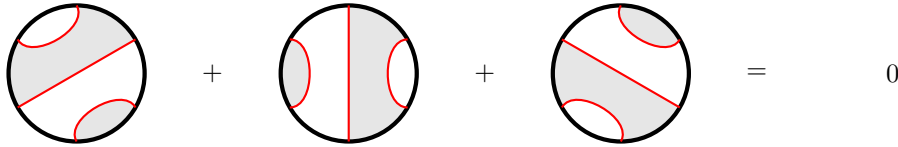
We can relate general chord diagrams, to chessboard chord diagrams, with the following observations.

If our "inner product" is supposed to be nondegenerate, then based on our previous theorem, we should set

$$\Gamma + \Gamma' + \Gamma'' = 0$$

for any triple of chord diagrams $\Gamma, \Gamma', \Gamma''$ obtained by bypass surgeries.

This relation is called the *bypass relation* and can be written schematically as



We therefore define a vector space V_n as follows: V_n is generated over \mathbb{Z}_2 by all chord diagrams with n chords, subject to the bypass relation, i.e. the relation that any three chord diagrams related by bypass surgeries sum to zero.

$$V_n = \frac{\mathbb{Z}_2 \langle \text{Chord diagrams with } n \text{ chords} \rangle}{\text{Bypass relation}}$$

Before taking a quotient, the vector space is freely generated by chord diagrams, and its dimension is equal to the number of chord diagrams with n chords, which is C_n , the n 'th Catalan number.

After the quotient, something interesting happens. The dimension is reduced to 2^{n-1} and a basis is rather familiar.

Theorem 4.4 ([30]). *The \mathbb{Z}_2 vector space V_n has dimension 2^{n-1} and the diagrams from chessboards of $n - 1$ squares form a basis.*

Indeed, there are 2^{n-1} configurations of pawns on a chessboard with $n-1$ squares, and these give the chessboard chord diagrams with n chords. (Each square, again, contributing one bit of information.) Alternatively, this is the \mathbb{Z}_2 vector space freely generated by words in p and q .

Corollary 4.5.

$$V_n \cong (p\mathbb{Z}_2 \oplus q\mathbb{Z}_2)^{\otimes(n-1)}.$$

There is much more structure in this vector space, encoding various combinatorial and representation-theoretic properties of chord diagrams. For instance, V_n has $2^{2^{n-1}}$ elements; the C_n of these which correspond to chord diagrams are distributed in a combinatorially interesting way.

To give an idea of how chord diagrams decompose into chessboard diagrams, we give a computation below.

$$\begin{aligned}
&= \Gamma_{ppqq} + \Gamma_{pqqp} + \Gamma_{qppq} + \Gamma_{qpqp} \\
&\cong ppqq + pqqp + qppq + qpqp
\end{aligned}$$

We should point out that this is not the only interesting basis of V_n . In [33] we defined a large class of bases, one from any *quadrangulation* of a surface.

5 Contact topology

As it turns out, all these curves on surfaces and stuffing into cylinders and 120° rotations describe 3-dimensional contact topology.

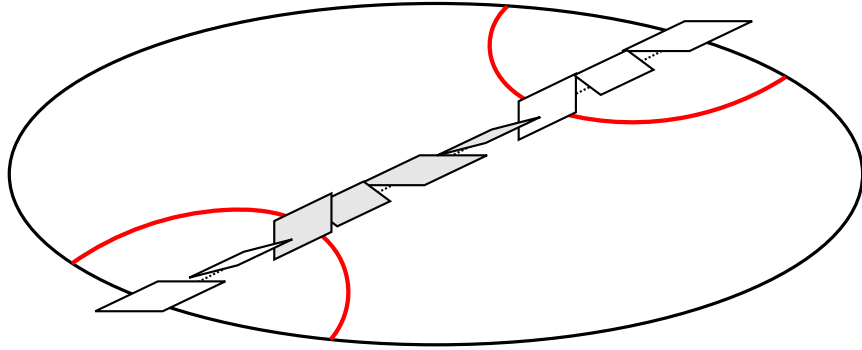
5.1 Chord diagrams and contact structures

First of all, a *chord diagram* Γ on a disc D describes a *contact structure* ξ_Γ on $D \times I$. This contact structure ξ_Γ consists of planes which are, roughly (and inaccurately) speaking,

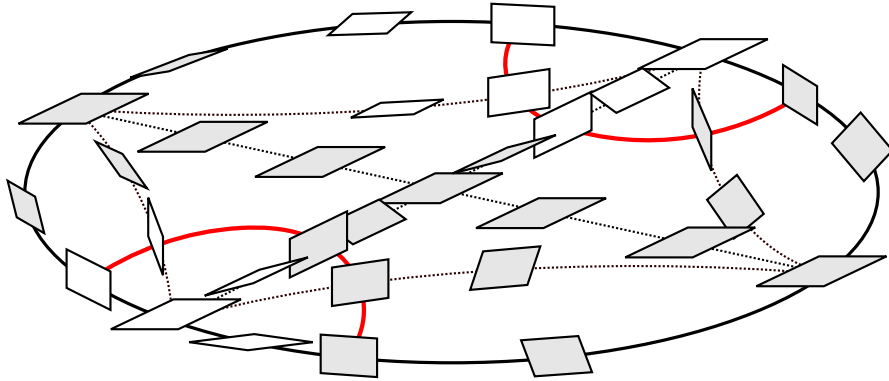
- Tangent to D around the boundary of D
- “Perpendicular” to D precisely along Γ .²

In the simple example below of a disc with two chords, as we proceed from one side to the other, crossing both chords, the contact planes rotate through a full 360° . A rough (and inaccurate) interpretation of the contact condition of non-integrability is that if we follow a curve C tangent to the contact structure ξ , dotted in the diagram, then the contact planes always rotate in the same direction about C .

²This statement makes no sense: a contact manifold has no metric, and no notion of perpendicularity! Rather, we should say that ξ is invariant under translations in the I direction, and the vector field in the I direction lies in ξ precisely along $\Gamma \times I$.



If we colour one side of the contact planes white and the other side black (grey) then the black and white regions in a chord diagram correspond to which side of the contact planes are visible from above.



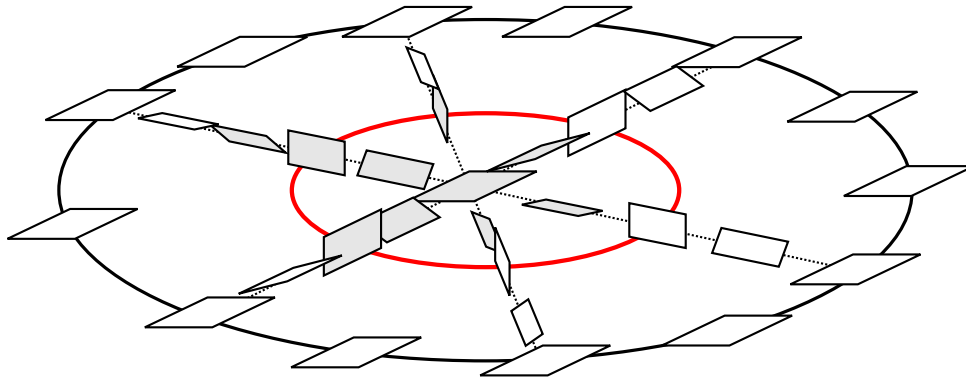
It turns out that, in a sense which can be made precise, there is only one “way up to isotopy” to draw contact planes consistent with the chords. The lines of intersection of the contact planes with the disc D trace out a foliation called the *characteristic foliation*, some curves of which are dotted in the diagrams. The characteristic foliation determines the germ of a contact structure near D . The characteristic foliation is transverse to the chords Γ (known as a *dividing set* in contact geometry) and in fact can be directed by a vector field to exponentially dilate an area form on each component of $D \setminus \Gamma$ exiting through Γ . All such foliations compatible with Γ give isotopic germs of contact structures. This is all part of the theory of *convex surfaces* developed by Giroux [16] in 1991.

5.2 Contact structures and overtwisted discs

Given a 3-manifold M , it’s an interesting and difficult question to find all the isotopy classes of contact structures on M .

Eliashberg in [7] showed that there are fundamentally two types of contact structures, called *overtwisted* and *tight*. An overtwisted contact structure is one that contains an object called an *overtwisted disc*; a tight contact structure is one that does not.

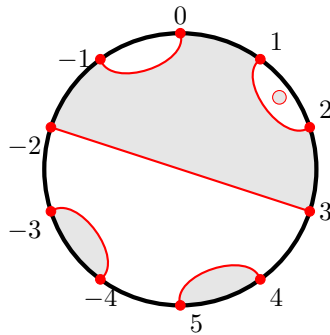
An *overtwisted disc* is a neighbourhood of a disc with a contact structure as shown. It corresponds to seeing a *contractible closed curve* in a chord diagram / dividing set on a disc.



Eliashberg in 1989 [7] proved that the classification of overtwisted contact structures on a 3-manifold M is equivalent to the homotopy classification of plane fields on M . In this case the contact geometry reduces to homotopy theory (which is well understood) and contact geometry offers nothing new. So contact topologists tend to regard an overtwisted disc as “spoiling” a contact structure or rendering it trivial. It is quite surprising that the presence of one particular disc in a contact 3-manifold has such global topological consequences.

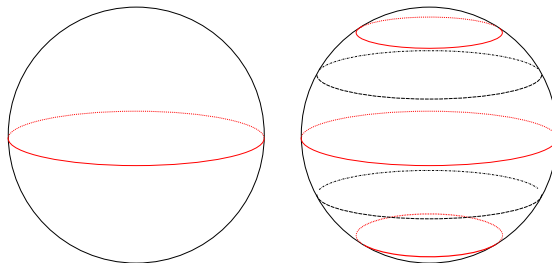
The *tight* contact structures on M , on the other hand, offer important information about the topology of M . Not every 3-manifold has a tight contact structure, as Etnyre and Honda proved in 1999 [11]; and the number of tight contact structures on a 3-manifold depends intricately on the topology of M (see, e.g., [17, 18, 21, 22]).

This makes it quite reasonable that we regard a diagram on a disc, with a red closed curve, as “trivial”, and count it as 0, such as in this figure seen previously:



5.3 Contact structures on spheres and balls

Contact structures in the neighbourhood of a sphere S^2 are again given by dividing sets on S^2 , which we continue to draw in red. It turns out that there is essentially only one tight contact structure in the neighbourhood of an S^2 , and that is given by a *single* connected curve. Any dividing set with more than one curve gives an overtwisted contact structure.



The sphere on the left has a tight contact neighbourhood. The sphere on the right has an overtwisted contact neighbourhood, and boundaries of two overtwisted discs are drawn in dashed black.

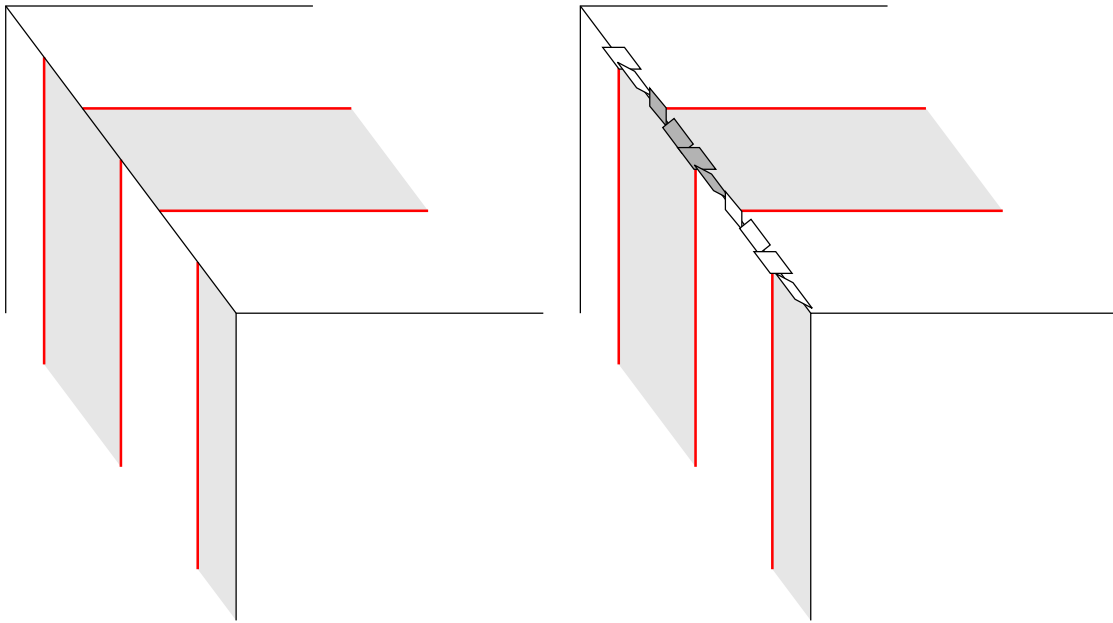
Now consider each sphere S^2 as the boundary of a ball B^3 . Contact structures on balls were classified by Eliashberg in 1992 [8]. Putting aside the overtwisted structures, there is essentially only one *tight* contact structure on the ball, up to isotopy. Given a dividing set on the boundary sphere which is connected, there is an essentially unique way to fill it in.

5.4 Contact corners

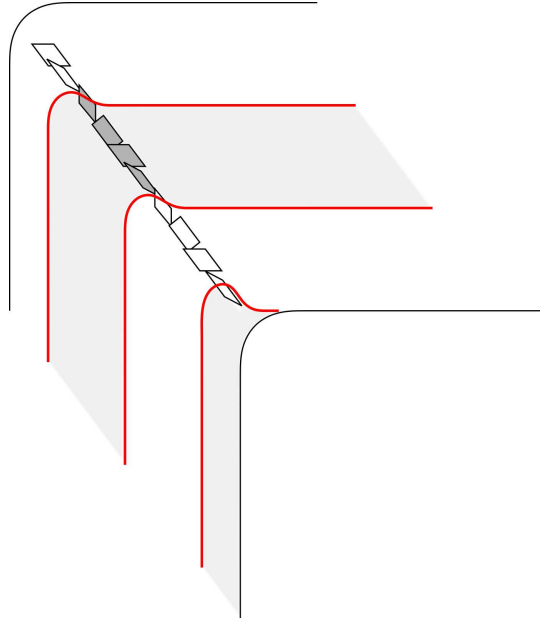
When considering a surface S with boundary ∂S in a contact 3-manifold (M, ξ) , we often require the contact planes to be tangent to the boundary ∂S . Above, all our diagrams of discs with contact planes have been drawn in this way.

When two surfaces meet along their boundary, forming a *corner*, then the dividing sets do not intersect along the corner — with our rough interpretation of dividing set as “where contact planes are perpendicular to the surface”, this would mean that the contact planes are perpendicular to both surfaces meeting at the corner, as well as tangent to the corner, which is impossible.

Rather, the contact planes rotate around the corner curve, and the the dividing sets *interleave* as shown.



It’s not too difficult to believe, then, that if we *round* the corners, then dividing curves behave as we proposed earlier.



It follows then that the “inner product” on chord diagrams, involving insertion into a cylinder, has a simple contact geometry interpretation.

Proposition 5.1. *Let Γ_0, Γ_1 be chord diagrams. The following are equivalent:*

- (i) $\langle \Gamma_0 | \Gamma_1 \rangle = 1$.
- (ii) *The solid cylinder with dividing set Γ_0 on the bottom and Γ_1 on the top has a tight contact structure.*

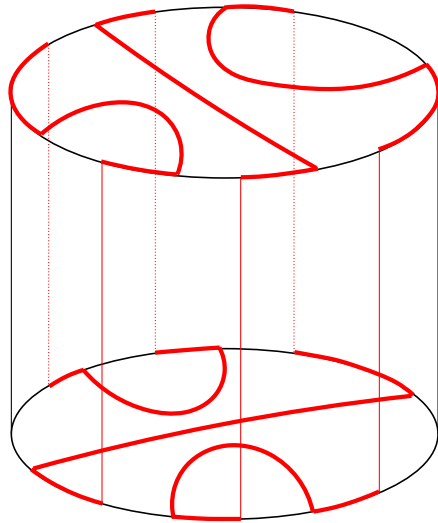
□

The “finger moves” we saw earlier for curves on the cylinder now correspond simply to isotopy of contact structures near the sphere, and in the ball.

5.5 Bypasses

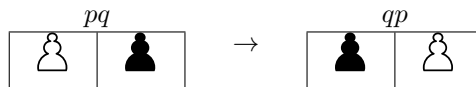
Suppose we start from a surface S with a dividing set on it — such as a disc with a chord diagram. This describes a contact structure near S ; effectively, on $S \times I$. We could then try to build up a contact manifold by gluing on more pieces on top of $S \times I$, and gluing them up.

It turns out that any contact 3-manifold can be built up in this way, using only fundamental building blocks called *bypasses*. Drawn in terms of chord diagrams and cylinders, a bypass is actually something we saw earlier in computing $\langle \Gamma_{pq} | \Gamma_{qp} \rangle$. It is drawn below.

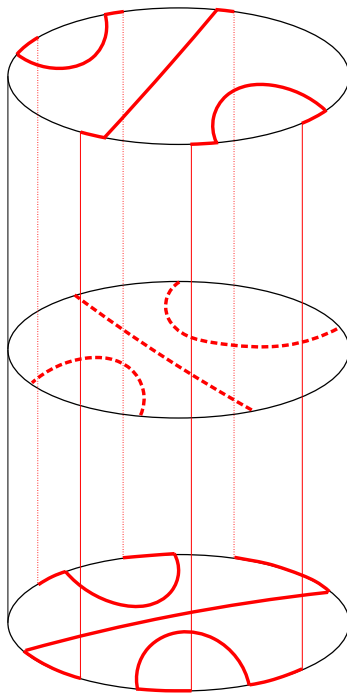


A bypass is nothing but a particular contact 3-ball. We consider the cylinder as bounding a 3-ball; from Eliashberg's theorem we know that the contact structure on the boundary of the cylinder extends uniquely throughout the ball to a tight contact structure.

Thus, the elementary building blocks of contact topology are, in a certain sense, precisely the elementary pawn moves



However, beware! If you stack two bypasses directly on top of each other, you will find an over-twisted disc. We might say that the shortest step in contact topology is half way to oblivion.



To build a tight contact structure, you'll need to place bypasses in more sophisticated locations than just on top of each other.

We see that addition of a bypass on top of a disc performs a 120° rotation in the chord diagram. This is precisely the operation described earlier and we defined the vector space V_n by setting triples of bypass-related chord diagrams to sum to zero.

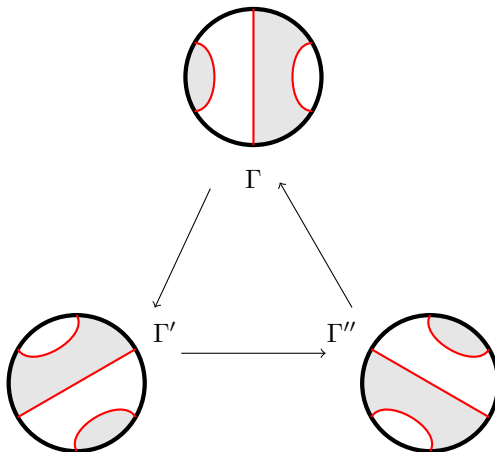
In fact, our creation and annihilation operators, which added or closed off curves in a chord diagram, can also be seen as building onto an existing contact structure.

5.6 The contact category

Honda has introduced the concept of the *contact category* [20]. For a disc, the contact category \mathcal{C}_n consists of objects and morphisms as follows.

- (i) The *objects* are chord diagrams with n chords, i.e. contact structures near discs D .
- (ii) The *morphisms* $\Gamma_0 \rightarrow \Gamma_1$ are contact structures on the cylinder $D \times I$, with Γ_0 on the bottom and Γ_1 on the top.
- (iii) *Composition of morphisms* is given by stacking cylinders on top of each other.

Honda showed that this category has many of the properties of a *triangulated category*. In particular bypass triples behave very much like exact triangles. If $\Gamma, \Gamma', \Gamma''$ are successively obtained from each other by adding bypasses, then the composition of any two is overtwisted, hence 0.



But we have seen here that chord diagrams themselves can be described by pawns and chessboards — and the dynamics of pawns on chessboards is itself essentially a *partial order*, which is a type of category. Our work in [30] defined a *contact 2-category* out of this geometric structure.

In fact, looking now at the vector space V_n we defined, we have essentially

$$V_n = \frac{\mathbb{Z}_2\langle \text{Chord diagrams with } n \text{ chords} \rangle}{\text{Bypass relation}} = \frac{\mathbb{Z}_2\langle \text{Objects of } \mathcal{C}_n \rangle}{\text{Exact triangles sum to zero}}$$

and thus V_n is the *Grothendieck group* of \mathcal{C}_n .

5.7 Combinatorics of contact geometry

To summarise, we’ve now seen various ways in which the combinatorics of chords on discs and cylinders, quantum pawn dynamics, and contact geometry are connected:

- (i) Any chord diagram describes a contact structure near a disc.
- (ii) The “inner product” $\langle \cdot | \cdot \rangle$ can be interpreted as pawn dynamics on a chessboard, insertion of chord diagrams into a cylinder, or the existence of tight contact structures in solid cylinders.

- (iii) Creation and annihilation operators can be defined as pawn/antipawn creation/annihilation, or the insertion/closure of curves in a chord diagram, or building a contact structure.
- (iv) The adjoint property of creation and annihilation can be interpreted via the combinatorics of chessboards, finger moves on cylinders, or isotopy of contact structures on the sphere.
- (v) A diagram on a disc with a closed curve describes an overtwisted contact structure, hence is “trivial” in terms of contact geometry. Similarly for curves on a cylinder with more than one component.
- (vi) Bypass triples of chord diagrams correspond to: building blocks of contact 3-manifolds; triples which must sum to zero for the “insertion into cylinder inner product” $\langle \cdot | \cdot \rangle$ to be nondegenerate; and exact triangles in the contact category.
- (vii) Chessboards correspond to a special class of chord diagrams, which look like slalom ski slopes, forming a basis for V_n , the Grothendieck group of the contact category of the disc.

Much of this structure is similar to *topological quantum field theories* (TQFT’s). In particular, we have assigned vector spaces V_n to discs, which are 2-dimensional. To cylinders, which we can think of as $(2+1)$ -dimensional, or as the “time evolution” of a disc, we have associated an element of \mathbb{Z}_2 , or a morphism in the contact category. And we have the “inner product” which describes not so much a probability amplitude, but a *possibility amplitude* for a tight contact structure to evolve from one disc to another. We can call this structure *contact TQFT* and summarise much of what we have said as

$$\text{Contact TQFT} \cong \text{QPD}.$$

6 Holomorphic invariants

In fact, all of this structure actually arose from some of the holomorphic invariants mentioned earlier.

6.1 Sutured Floer homology

Sutured Floer homology is a variant of Heegaard Floer homology. A sutured 3-manifold (M, Γ) is (roughly) a 3-manifold M with boundary, and some curves Γ on the boundary ∂M , which divide ∂M into positive and negative regions. Sutured 3-manifolds were studied by Gabai in the 1980s in the context of foliation theory [12], but also describe contact 3-manifolds with boundary — indeed, the diagrams we have drawn with discs, red curves, and complementary regions coloured black and white are sutures. There are close relationships between foliation theory, sutured manifolds and contact topology [23, 9].

Sutured Floer homology, again speaking roughly and imprecisely, is defined as follows. A sutured 3-manifold³ (M, Γ) has a Heegaard decomposition, consisting of a Heegaard surface with boundary Σ , and two sets of curves $\alpha_1, \dots, \alpha_k$ and β_1, \dots, β_k on Σ . The manifold can be recovered from the decomposition by gluing to $\Sigma \times [0, 1]$ discs along $\alpha_i \times \{0\}$ and $\beta_i \times \{1\}$. The sutures lie along $\partial \Sigma \times [0, 1]$, the negative part of the boundary is $\Sigma \times \{0\}$ surgered along the $\alpha_i \times \{0\}$, and the positive part of the boundary is $\Sigma \times \{1\}$ surgered along the $\beta_i \times \{1\}$. We then consider holomorphic curves

$$u : S \longrightarrow \Sigma \times I \times \mathbb{R}$$

where S is a Riemann surface with boundary and with negative boundary punctures p_1, \dots, p_k and positive boundary punctures q_1, \dots, q_k , satisfying conditions including the following (among others)⁴:

³Strictly speaking, a *balanced* sutured 3-manifold.

⁴There are also some more technical conditions including conditions on the complex structure, non-constant preprojections of u , and finite energy. We just highlight the fact that the boundary conditions come from 3-manifold topology. See [27] and [26].

(i) u sends ∂S to

$$((\cup_i \alpha_i) \times \{1\} \times \mathbb{R}) \cup ((\cup_i \beta_i) \times \{0\} \times \mathbb{R}).$$

(ii) u asymptotically sends each p_i to $-\infty$ in the \mathbb{R} coordinate, and each q_i to $+\infty$.

(iii) Each $u^{-1}(\alpha_i \times \{1\} \times \mathbb{R})$ and $u^{-1}(\beta_i \times \{0\} \times \mathbb{R})$ consists of precisely one segment of $\partial S \setminus \{p_1, \dots, p_k, q_1, \dots, q_k\}$.

The point I want to make here is that, whatever the precise boundary conditions for holomorphic curves, they are defined in terms of the Heegaard decomposition. At $+\infty$, the curve $u(S)$ (or rather, its projection to Σ) runs through several intersections of α and β curves. Indeed, at $+\infty$ we obtain a set of points $z_1, \dots, z_k \in \Sigma$, such that

$$z_1 \in \alpha_1 \cap \beta_{\sigma(1)}, \quad z_2 \in \alpha_2 \cap \beta_{\sigma(2)}, \quad \dots, \quad z_k \in \alpha_k \cap \beta_{\sigma(k)}$$

for some permutation σ . We obtain a similar set of intersection points at $-\infty$.

We define a chain complex \widehat{CF} generated over \mathbb{Z}_2 by sets of such intersection points,

$$\mathbf{x} = \{z_1, \dots, z_k\}.$$

If we fix \mathbf{x} and \mathbf{y} to be two sets of intersection points, then we consider the moduli space $\mathcal{M}(\mathbf{x}, \mathbf{y})$ with boundary conditions given by \mathbf{x} at $+\infty$ and \mathbf{y} at $-\infty$. The dimension of this moduli space can be given in terms of the topology of \mathbf{x} and \mathbf{y} (an *index formula*). It follows, with some serious analysis, that we can define a differential to count index-1 curves

$$\partial \mathbf{x} = \sum_{\dim \mathcal{M}(\mathbf{x}, \mathbf{y})=1} \# \widehat{\mathcal{M}}(\mathbf{x}, \mathbf{y}) \cdot \mathbf{y}$$

and $\partial^2 = 0$. (Here $\widehat{\mathcal{M}}$ is the quotient of the moduli space \mathcal{M} by the action of translation along \mathbb{R} ; so $\widehat{\mathcal{M}}$ consists of finitely many points.) Then we may take the *homology* and Ozsváth–Szabó (for closed manifolds [40, 39]) and Juhasz (for sutured manifolds [26]) showed that this homology is *independent of the choice of Heegaard decomposition* (and independent of other technical choices made along the way). That is, the resulting homology is a sutured manifold invariant which we may call $SFH(M, \Gamma)$.

Sutured Floer homology also has the nice property that any contact structure ξ on (M, Γ) gives a *contact element* $c(\xi)$ in $SFH(M, \Gamma)$ [24].⁵

It turns out that our subject matter in this note has been sutured Floer homology. In particular, SFH of solid tori $D^2 \times S^1$. If we let F_n consist of $2n$ points on the boundary ∂D^2 , so that $F_n \times S^1$ forms a set of sutures on $D^2 \times S^1$, then we have the following.

Theorem 6.1 ([30]).

$$V_n \cong SFH(D^2 \times S^1, F_n \times S^1).$$

Moreover, any chord diagram Γ in V_n corresponds to a tight contact structure ξ_Γ on $(D^2 \times S^1, F_n \times S^1)$, unique up to isotopy, and the isomorphism takes $\Gamma \mapsto c(\xi_\Gamma)$.

Therefore, all of the structure we have been discussing, is contained in sutured Floer homology of one of the simplest sutured 3-manifolds, namely a solid torus.

6.2 Embedded contact homology and string homology

Another holomorphic invariant is *embedded contact homology*, defined by Hutchings [25]. Given a 3-manifold M , we consider a contact structure ξ and a contact form α . From the 1-form α there is a natural vector field called the *Reeb vector field* R , which is defined by

$$d\alpha(R, \cdot) = 0, \quad \alpha(R) = 1.$$

⁵This element is only defined up to sign; however if we take \mathbb{Z}_2 coefficients, no ambiguity arises.

Embedded contact homology counts holomorphic curves in the *symplectization* $M \times \mathbb{R}$ with symplectic form $d(e^t \alpha)$, where t is the coordinate on \mathbb{R} . The most important condition prescribed on holomorphic curves is that the curves must approach *Reeb orbits*, i.e. orbits of R , as $t \rightarrow \pm\infty$. Roughly, given certain collections of Reeb orbits

$$\gamma^+ = \{\gamma_1^+, \dots, \gamma_n^+\} \quad \text{and} \quad \gamma^- = \{\gamma_1^-, \dots, \gamma_n^-\},$$

(some of which might be repeated and some of which might be covered multiple times), we consider the moduli space $\mathcal{M}(\gamma^+, \gamma^-)$ of *embedded* holomorphic curves which approach γ^+ at $+\infty$ and γ^- at $-\infty$. Again there is an index formula giving the dimension of the moduli space in terms of the contact geometry of the γ_i^\pm and again, after a lot of work, it is possible to define a differential

$$\partial \gamma^+ = \sum_{\dim \mathcal{M}(\gamma^+, \gamma^-)=1} \# \widehat{\mathcal{M}}(\gamma^+, \gamma^-) \cdot \gamma^-$$

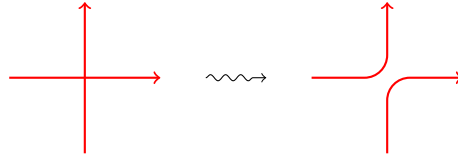
where $\partial^2 = 0$. The homology of this complex is *embedded contact homology*, *ECH*. It's possible to define *ECH* also for sutured manifolds [6].

It has recently been shown that embedded contact homology is isomorphic to Heegaard Floer homology (see e.g. [5]). Note that *ECH*, even though it is defined in terms of a contact form, turns out not to depend on the contact structure at all, but is a smooth manifold invariant.

This deep isomorphism between *ECH* and *HF* implies that we should be able to obtain all of the combinatorial and algebraic structure we have discussed above, from embedded contact homology. Based on work of Cieliebak–Latschev [4], as discussed in [34], one is led to consider the following ideas.

Take a disc D with $2n$ points F marked on the boundary, as we have seen with chord diagrams. But now consider sets of curves on D with boundary F , which are not necessarily chord diagrams — the curves may intersect. We call these *string diagrams*.

We define a \mathbb{Z}_2 vector space freely generated by string diagrams, up to homotopy relative to boundary. On this vector space, there is a *differential* ∂ defined by *resolving crossings*. That is, given a string diagram s , ∂s is the sum of string diagrams, each obtained by resolving one of the crossings of s as shown.

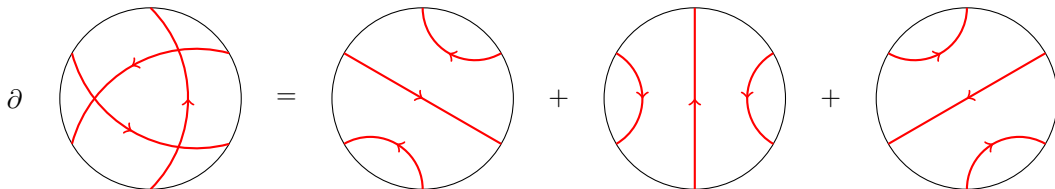


It turns out that $\partial^2 = 0$, and the homology is something with which we are by now familiar.

Theorem 6.2 ([34]).

$$HS \cong V_n \cong \frac{\mathbb{Z}_2 \langle \text{chord diagrams on } (D^2, F) \rangle}{\text{Bypass relation}}$$

The “reason” for this relation — which is far from a proof — is the following picture.



This idea of taking vector spaces generated by topological classes of curves, and resolving their crossings, is closer to the sorts of objects studied in *string topology* (e.g. [3]) and as such we can say that contact topology and sutured Floer homology are expressed as a *string homology*. Further details and generalisations of this result are given in [34].

References

- [1] V. I. Arnold, *Symplectic geometry and topology*, J. Math. Phys. **41** (2000), no. 6, 3307–3343. MR 1768639 (2001h:53108)
- [2] Frédéric Bourgeois, *Introduction to contact homology*, unpublished notes, <http://homepages.ulb.ac.be/~fbourgeo/pspdf/Berder.pdf>, 2003.
- [3] Moira Chas and Dennis Sullivan, *String topology*, <http://arxiv.org/abs/math/9911159>, 2008.
- [4] Kai Cieliebak and Janko Latschev, *The role of string topology in symplectic field theory*, New perspectives and challenges in symplectic field theory, CRM Proc. Lecture Notes, vol. 49, Amer. Math. Soc., Providence, RI, 2009, pp. 113–146. MR 2555935 (2010j:53187)
- [5] Vincent Colin, Paolo Ghiggini, and Ko Honda, *Equivalence of Heegaard Floer homology and embedded contact homology via open book decompositions*, Proc. Natl. Acad. Sci. USA **108** (2011), no. 20, 8100–8105. MR 2806645 (2012f:53190)
- [6] Vincent Colin, Paolo Ghiggini, Ko Honda, and Michael Hutchings, *Sutures and contact homology I*, Geom. Topol. **15** (2011), no. 3, 1749–1842. MR 2851076
- [7] Y. Eliashberg, *Classification of overtwisted contact structures on 3-manifolds*, Invent. Math. **98** (1989), no. 3, 623–637. MR MR1022310 (90k:53064)
- [8] Yakov Eliashberg, *Contact 3-manifolds twenty years since J. Martinet’s work*, Ann. Inst. Fourier (Grenoble) **42** (1992), no. 1-2, 165–192. MR MR1162559 (93k:57029)
- [9] Yakov M. Eliashberg and William P. Thurston, *Confoliations*, University Lecture Series, vol. 13, American Mathematical Society, Providence, RI, 1998. MR 1483314 (98m:53042)
- [10] John B. Etnyre, *Introductory lectures on contact geometry*, Topology and geometry of manifolds (Athens, GA, 2001), Proc. Sympos. Pure Math., vol. 71, Amer. Math. Soc., Providence, RI, 2003, pp. 81–107. MR 2024631 (2005b:53139)
- [11] John B. Etnyre and Ko Honda, *On the nonexistence of tight contact structures*, Ann. of Math. (2) **153** (2001), no. 3, 749–766. MR 1836287 (2002d:53119)
- [12] David Gabai, *Foliations and the topology of 3-manifolds*, Bull. Amer. Math. Soc. (N.S.) **8** (1983), no. 1, 77–80. MR 682826 (84h:57014)
- [13] Hansjörg Geiges, *A brief history of contact geometry and topology*, Expo. Math. **19** (2001), no. 1, 25–53. MR 1820126 (2002c:53129)
- [14] ———, *Christiaan Huygens and contact geometry*, Nieuw Arch. Wiskd. (5) **6** (2005), no. 2, 117–123. MR 2149304
- [15] ———, *An introduction to contact topology*, Cambridge Studies in Advanced Mathematics, vol. 109, Cambridge University Press, Cambridge, 2008. MR 2397738 (2008m:57064)
- [16] Emmanuel Giroux, *Convexité en topologie de contact*, Comment. Math. Helv. **66** (1991), no. 4, 637–677. MR MR1129802 (93b:57029)
- [17] ———, *Structures de contact en dimension trois et bifurcations des feuilletages de surfaces*, Invent. Math. **141** (2000), no. 3, 615–689. MR MR1779622 (2001i:53147)
- [18] ———, *Structures de contact sur les variétés fibrées en cercles au-dessus d’une surface (contact structures on manifolds that are circle-bundles over a surface)*, Comment. Math. Helv. **76** (2001), no. 2, 218–262. MR MR1839346 (2002c:53138)
- [19] M. Gromov, *Pseudoholomorphic curves in symplectic manifolds*, Invent. Math. **82** (1985), no. 2, 307–347. MR MR809718 (87j:53053)
- [20] Ko Honda, *Contact structures, Heegaard Floer homology and triangulated categories*, in preparation.
- [21] ———, *On the classification of tight contact structures. I*, Geom. Topol. **4** (2000), 309–368 (electronic). MR MR1786111 (2001i:53148)
- [22] ———, *On the classification of tight contact structures. II*, J. Differential Geom. **55** (2000), no. 1, 83–143. MR MR1849027 (2002g:53155)
- [23] Ko Honda, William H. Kazez, and Gordana Matić, *Convex decomposition theory*, Int. Math. Res. Not. (2002), no. 2, 55–88. MR MR1874319 (2002k:57064)

- [24] Ko Honda, William H. Kazez, and Gordana Matić, *The contact invariant in sutured Floer homology*, *Invent. Math.* **176** (2009), no. 3, 637–676. MR 2501299 (2010g:57037)
- [25] Michael Hutchings, *An index inequality for embedded pseudoholomorphic curves in symplectizations*, *J. Eur. Math. Soc. (JEMS)* **4** (2002), no. 4, 313–361. MR 1941088 (2004b:53148)
- [26] András Juhász, *Holomorphic discs and sutured manifolds*, *Algebr. Geom. Topol.* **6** (2006), 1429–1457 (electronic). MR MR2253454 (2007g:57024)
- [27] Robert Lipshitz, *A cylindrical reformulation of Heegaard Floer homology*, *Geom. Topol.* **10** (2006), 955–1097 (electronic). MR MR2240908 (2007h:57040)
- [28] Robert Lipshitz, Peter Ozsvath, and Dylan Thurston, *Bordered Heegaard Floer homology: Invariance and pairing*, <http://arxiv.org/abs/0810.0687>, 2008.
- [29] Ciprian Manolescu, Peter Ozsváth, and Sucharit Sarkar, *A combinatorial description of knot Floer homology*, *Ann. of Math. (2)* **169** (2009), no. 2, 633–660. MR MR2480614 (2009k:57047)
- [30] Daniel V. Mathews, *Chord diagrams, contact-topological quantum field theory, and contact categories*, *Algebraic & Geometric Topology* **10** (2010), no. 4, 2091–2189.
- [31] ———, *Sutured Floer homology, sutured TQFT and noncommutative QFT*, *Algebr. Geom. Topol.* **11** (2011), no. 5, 2681–2739. MR 2846909 (2012m:57057)
- [32] ———, *Sutured TQFT, torsion, and tori*, <http://arxiv.org/abs/1102.3450>, submitted for publication, 2011.
- [33] ———, *Itsy bitsy topological field theory*, <http://arxiv.org/abs/1201.4584>, 2012.
- [34] Daniel V. Mathews and Eric Schoenfeld, *Dimensionally-reduced sutured floer homology as a string homology*, <http://arxiv.org/abs/1210.7394>, 2012.
- [35] Dusa McDuff and Dietmar Salamon, *Introduction to symplectic topology*, second ed., Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 1998. MR MR1698616 (2000g:53098)
- [36] ———, *J-holomorphic curves and symplectic topology*, American Mathematical Society Colloquium Publications, vol. 52, American Mathematical Society, Providence, RI, 2004. MR 2045629 (2004m:53154)
- [37] Peter Ozsváth and Zoltán Szabó, *Holomorphic disks and genus bounds*, *Geom. Topol.* **8** (2004), 311–334. MR 2023281 (2004m:57024)
- [38] ———, *Holomorphic disks and knot invariants*, *Adv. Math.* **186** (2004), no. 1, 58–116. MR 2065507 (2005e:57044)
- [39] ———, *Holomorphic disks and three-manifold invariants: properties and applications*, *Ann. of Math. (2)* **159** (2004), no. 3, 1159–1245. MR MR2113020 (2006b:57017)
- [40] ———, *Holomorphic disks and topological invariants for closed three-manifolds*, *Ann. of Math. (2)* **159** (2004), no. 3, 1027–1158. MR MR2113019 (2006b:57016)
- [41] ———, *Heegaard diagrams and Floer homology*, *International Congress of Mathematicians. Vol. II*, Eur. Math. Soc., Zürich, 2006, pp. 1083–1099. MR MR2275636 (2008h:57048)
- [42] Hermann Weyl, *The Classical Groups. Their Invariants and Representations*, Princeton University Press, Princeton, N.J., 1939. MR 0000255 (1,42c)
- [43] John Archibald Wheeler, *Information, physics, quantum: the search for links*, *Foundations of quantum mechanics in the light of new technology (Tokyo, 1989)*, Phys. Soc. Japan, Tokyo, 1990, pp. 354–368. MR 1105973