Chapter 1

Prologue

The main actors of this book are **classical theta functions** and the **Gaussian linking number**. They were brought together by Edward Witten using a quantum field theory whose Lagrangian is the Chern-Simons functional. In this chapter we offer the reader a first encounter with theta functions, the linking number, and Witten's Chern-Simons theory. The discussion is less formal, less detailed, and less rigorous then the rest of the book, and should be read like a historical summary. We will revisit these notions later, with more rigor and detail.

1.1 The history of theta functions

In the development of mathematics, theta functions appeared in early 19th century, as tools for studying elliptic integrals and elliptic functions. The reader should be aware that a great body of mathematics was devoted to elliptic integrals and functions. Below we only sketch a few of the most important contributions; for more details we recommend [Houzel (1978)].

1.1.1 *Elliptic integrals and theta functions*

Elliptic integrals appeared in the 17th and 18th century in computations of arc-lengths of curves and in models of classical mechanics such as pendulums and springs. The first example was the integral computing the arc-length of an **ellipse**, which gave the name of the entire class. For the ellipse

$$x = a\sin\phi, \quad y = b\cos\phi, \quad \phi \in [0, 2\pi]$$

the arc-length is computed by the integral

$$\int_P^Q ds = a\int_{\phi_P}^{\phi_Q} \sqrt{1-k^2\sin^2\phi d\phi}$$

where $k^2 = 1-b^2/a^2$ is the eccentricity. The integrand is not the derivative of an elementary function.

An example from mechanics is the integral

$$y = \int \frac{x^2 dx}{\sqrt{1-x^4}},$$

which Jacob Bernoulli found to describe the equilibrium position of an elastic rod with fixed extremities, subject to a force. Neither is this integral computable through elementary functions.

Lagrange was the first to consider the general case. For him elliptic integrals are of the form

$$\int R(x,y)dx$$

where R is a rational function, and $y = \sqrt{P(x)}$ with P a polynomial of degree 3 or 4 without multiple roots. Legendre reduced elliptic integrals to three types:

$$\int \frac{d\phi}{\sqrt{1-k^2\sin^2\phi}}, \int \sqrt{1-k^2\sin^2\phi}d\phi, \int \frac{d\phi}{(1+n\sin^2\phi)\sqrt{1-k^2\sin^2\phi}},$$

where $x = \sin\phi$.

Since these integrals cannot be reduced to known functions, an intense effort was devoted to their study and their approximate computation. An important role was played by **addition formulas**, first discovered by Euler in the case of the arc-length of the lemniscate, which found their apogee in the works of Abel and Jacobi.

To explain these formulas, let us take a look at the simpler situation of the trigonometric functions $\sin u$ and $\cos u$. We are familiar with them because mathematics started with geometry. But say if mathematics had started with polynomials, then the natural introduction of sine and cosine would have been through their inverse functions

$$\arcsin u = \int_0^u \frac{dx}{\sqrt{1-x^2}}, \quad \arccos u = -\int_0^u \frac{dx}{\sqrt{1-x^2}}.$$

Then $\sin u$ and $\cos u$ would be the inverses of these integrals.

One has the addition formula

$$\sin(u+v) = \sin u\cos v + \cos u \sin v.$$

Moreover, $\sin u$ and $\cos u$ are periodic functions of period 2π. This period is 4 times the "complete integral"

$$\arcsin(1) - \arcsin(0) = \int_0^1 \frac{dx}{\sqrt{1-x^2}}.$$

Abel had the idea of considering the inverse functions of elliptic integrals. These are what we now call **elliptic functions**. Jacobi, inspired by Abel, defined the inverse $\operatorname{am} u$ of the elliptic integral of the first kind

$$u = \int \frac{d\phi}{\sqrt{1-k^2\sin^2\phi}}.$$

Gudermann later introduced the notation

$$\operatorname{sn} u = \sin \operatorname{am} u, \quad \operatorname{cn} u = \cos \operatorname{am} u, \quad \operatorname{dn} u = \sqrt{1-k^2\sin^2 \operatorname{am} u}.$$

The functions sn and cn are periodic with period $4K$ and dn is periodic with period $2K$, where K is the complete integral

$$u(\pi/2) - u(0) = \int_0^{\frac{\pi}{2}} \frac{d\phi}{\sqrt{1-k^2\sin^2\phi}}.$$

Abel showed that elliptic functions satisfy addition theorems similar to those for trigonometric functions, for example

$$\operatorname{sn}(u+v) = \frac{\operatorname{sn} u \operatorname{cn} v \operatorname{dn} v + \operatorname{sn} v \operatorname{cn} u \operatorname{dn} u}{1-k^2\operatorname{sn}^2 u \operatorname{sn}^2 v}.$$

Jacobi was able to extend these elliptic functions to a complex variable. In this context they differ from the trigonometric functions sine and cosine by two significant properties: they are meromorphic and they are doubly periodic. The periods of sn, cn, dn are respectively $(4K, 2iK')$, $(4K, 2(K+iK'))$, $(2K, 4iK')$, where

$$K' = \int_0^{\pi/2} \frac{d\phi}{\sqrt{1-k'^2\sin^2\phi}}, \text{ with } k' = \sqrt{1-k^2}.$$

Jacobi gave these three elliptic functions global analytical representations as quotients of holomorphic functions

$$\operatorname{sn} u = \frac{1}{\sqrt{k}}\frac{H(u)}{\Theta(u)}, \quad \operatorname{cn} u = \sqrt{\frac{k'}{k}}\frac{H_1(u)}{\Theta(u)}, \quad \operatorname{dn} u = \sqrt{k'}\frac{\Theta_1(u)}{\Theta(u)},$$

by introducing the four **theta functions**

$$\Theta\left(\frac{2Kx}{\pi}\right) = \vartheta(x;q) = \sum_{n\in\mathbb{Z}}(-1)^n q^{n^2} e^{2nix},$$

$$H\left(\frac{2Kx}{\pi}\right) = \vartheta_1(x;q) = \sum_{n\in\mathbb{Z}} i^{2n+1} q^{(n+1/2)^2} e^{(2n+1)ix},$$

$$H_1\left(\frac{2Kx}{\pi}\right) = \vartheta_2(x;q) = \sum_{n\in\mathbb{Z}} q^{(n+1/2)^2} e^{(2n+1)ix}$$

$$\Theta_1\left(\frac{2Kx}{\pi}\right) = \vartheta_3(x;q) = \sum_{n\in\mathbb{Z}} q^{n^2} e^{2nix},$$

where $q = e^{-\pi K'/K}$. Θ_1 and H_1 are computed in terms of Θ and H as $\Theta_1(u) = \Theta(K-u)$ and $H_1(u) = H(K-u)$.

The functions Θ and H are not doubly periodic, as Liouville's theorem would prohibit that, nevertheless they are periodic with periods $2K$ respectively $4K$ and they are as close as being doubly periodic as possible with

$$\Theta(u+2iK') = -e^{-i\pi u} q^{-1}\Theta(u),$$
$$H(u+2iK') = -e^{-\pi K'/K + 2\pi K' - i\pi u} q^{-1} H(u).$$

Nowadays it is customary to replace x by πz, and to use the parameter τ instead of q, with $q = \exp(\pi i\tau)$, allowing τ to range in the upper half-plane of the complex plane (so $K'/K = -i\tau$). Also, following Riemann, one emphasizes the theta function

$$\theta(z;\tau) = \theta_{00}(z;\tau) = \vartheta_3(x;q) = \sum_{n\in\mathbb{Z}} e^{\pi i\tau n^2 + 2\pi inz},$$

and relates the others to it by

$$\theta_{01}(z;\tau) = \vartheta(x;q) = \theta\left(z+\frac{1}{2};\tau\right),$$

$$\theta_{10}(z;\tau) = \vartheta_2(x;q) = e^{\frac{1}{4}\pi i\tau + \pi iz}\theta\left(z+\frac{1}{2}\tau;\tau\right)$$

$$\theta_{11}(z;\tau) = -\vartheta_1(x;q) = e^{\frac{1}{4}\pi i\tau + \pi i(z+\frac{1}{2})}\theta\left(z+\frac{1}{2}\tau+\frac{1}{2};\tau\right).$$

The periodicity condition reads

$$\theta(z+m+n\tau;\tau) = e^{-\pi in^2\tau - 2\pi inz}\theta^{\tau}(z;\tau).$$

Changes of coordinates were employed in order to improve the numerical computations of elliptic integrals. They led to the study of the behavior of

elliptic functions and of theta functions under coordinate transformations. Jacobi discovered the identity

$$\theta\left(\frac{z}{\tau}; -\frac{1}{\tau}\right) = (-i\tau)^{1/2} \exp\left(\frac{\pi}{\tau} i z^2\right) \theta(z;\tau).$$

This has led to the discovery of the action of the entire modular group

$$PSL(2,\mathbb{Z}) = \left\{ z \mapsto \frac{az+b}{cz+d} \,|\, ad - bc - 1 \right\}$$

on theta functions.

Theta functions were also used previously by Gauss for the same purpose as Jacobi, in his work on the arithmetic-geometric mean which he related to the elliptic integral of the first kind. They were used by Jacob Bernoulli and Euler in number theory, and by Poisson and Fourier for solving the heat equation. For the latter purpose one restricts the variable z to a real number x and the parameter τ to it with t real. Then

$$\frac{\partial}{\partial t}\theta(x;it) = \frac{1}{4\pi}\frac{\partial^2}{\partial x^2}\theta(x;it),$$

with the initial condition $\theta(x;0) = \sum_{n\in\mathbb{Z}} \delta(x-n)$, where δ is Dirac's delta function.

1.1.2 *The work of Riemann*

Riemann introduced a completely new point of view in the theory of elliptic integrals and theta functions.

To study a function $w(z)$ defined by a polynomial equation $F(z,w) = 0$ of degree m in z and n in w, Riemann introduced a complex surface Σ (a Riemann surface), which is an n-sheeted covering of the plane, and on which w is a well defined holomorphic function. Adding the point at infinity turns Σ into a compact, orientable surface. Riemann's programme was to study the integrals of rational functions $R(z,w)$ along paths in Σ.

The Riemann surface associated to $w(z)$ can have the genus g equal to 0 (when w can be solved explicitly in terms of z, which case is totally uninteresting), 1 (the case studied extensively by Abel and Jacobi), or higher. The function

$$u(x) = \int_a^x R(z(t), w(t))dt$$

is univalent in the complement of a finite collection of curves that form a basis for the first homology group of Σ. The $2g$ integrals of $R(z,w)$ along

these curves define the **periods** of the inverse function of u. The periods define a lattice in $\mathbb{C}^g$. The quotient of $\mathbb{C}$ by this lattice is the **Jacobian variety**. The function u defined by integrating $R(z,w)$ along paths in Σ is well defined from Σ to the Jacobian variety.

To study the elliptic functions on the Jacobian variety, Riemann introduced his multivariable theta function

$$\theta(\mathbf{z};\Pi) = \sum_{\mathbf{n}\in\mathbb{Z}^g} e^{\pi i \mathbf{n}^T \Pi \mathbf{n} + 2\pi i \mathbf{n}^T \mathbf{z}},$$

where Π is a $g \times g$ matrix with positive-definite imaginary part. The matrix Π depends on the complex structure of the Riemann surface Σ. In genus 1, Π is the complex number τ encountered above.

Example 1.1. These ideas are best understood when applied to the classical example of the Weierstrass elliptic integral

$$u(z) = \int_z^\infty \frac{dt}{\sqrt{t(t-1)(t-\lambda)}}, \quad 0 < \lambda < 1.$$

Here the Riemann surface is associated to the function w defined by $w^2 = \sqrt{z(z-1)(z-\lambda)}$. The function

$$w : \mathbb{C} \to \mathbb{C}, \quad w = \sqrt{z(z-1)(z-\lambda)}$$

is univalent on $\mathbb{C}\backslash([0,\lambda] \cup [1,\infty))$, so the Riemann surface associated to it is obtained by cutting open the complex sphere $\mathbb{C}P^1 = \mathbb{C} \cup \infty$ along two segments, $[0,\lambda]$ and $[1,\infty]$, taking two copies of the result, and gluing them along their boundaries as shown in Figure 1.1. This construction identifies the torus $S^1 \times S^1$ with the algebraic variety $z_1^2 z_2 = z_3(z_3 - z_2)(z_3 - \lambda z_2)$ in the complex projective space $\mathbb{C}P^2$ (with $w = z_1/z_2$, and $z = z_3$).

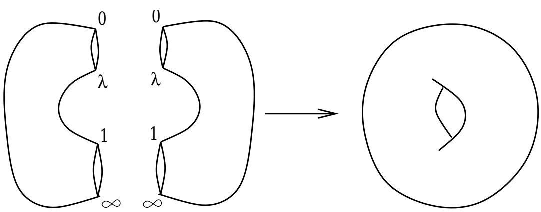


Fig. 1.1 The Riemann surface of $w = \sqrt{z(z-1)(z-\lambda)}$

The integral $u(z)$ is well defined only up to multiples of the values of two integrals. One of them is the integral along either of the branch cuts, the other is the integral along a curve that crosses each of the two branch

cuts once. These are **Riemann's periods** ω_1 and ω_2. Then u has well-defined values in the 2-dimensional torus $\mathbb{C}/(\mathbb{Z}\omega_1 + \mathbb{Z}\omega_2)$. This torus is the Jacobian variety associated to the elliptic integral.

The inverse of u is the **Weierstrass elliptic function** $\wp$, which can be interpreted either as a meromorphic function on the Jacobian variety, or as the doubly periodic meromorphic function on $\mathbb{C}$,

$$\wp(z;\omega_1,\omega_2) = \frac{1}{z^2} + \sum_{(m,n)\neq(0,0)} \left[\frac{1}{(z+m\omega_1+n\omega_2)^2} - \frac{1}{(m\omega_1+n\omega_2)^2} \right].$$

One can always arrange, by a change of coordinates, for ω_1 to be 1, in which case it is standard to write $\tau = \omega_2$. In terms of theta functions

$$\wp(z;\tau) = \pi^2\theta^2(0;\tau)\theta_{10}^2(0;\tau)\frac{\theta_{01}^2(z;\tau)}{\theta_{11}(z;\tau)} - \frac{1}{3}\pi^2(\theta^4(0;\tau)+\theta_{10}^4(0;\tau)).$$

In Chapter 4 we will develop the theory of Riemann's theta functions. For this we will employ the tools of quantum mechanics, since theta functions are yet another example of pure mathematics into which quantum theory offers new insights. We are motivated to take this approach by Witten's abelian Chern-Simons theory which is physical in its nature. The functions considered there are of a slightly finer structure. They include a parameter, which is an integer N that plays the role of Planck's constant; in Riemann's situation $N = 1$. They are called **theta functions with characteristics** in [Mumford (1983)].

The reader will notice that in Chapter 4 and the subsequent chapters we write $\theta^\tau(z)$ and $\theta^\Pi(\mathbf{z})$ instead of $\theta(z;\tau)$ and $\theta(\mathbf{z};\Pi)$. This is done so as to simplify formulas and computations, and to emphasize that z is the actual variable while τ and Π are (fixed) parameters depending on the complex structure of the Riemann surface.

1.2 The linking number

1.2.1 *The definition of the linking number*

In mathematics, a **knot** is an embedding of a circle in $\mathbb{R}^3$. The embedding of a disjoint union of finitely many circles in $\mathbb{R}^3$ is called a **link**; each circle is called a **link component**. Two knots or links are the same if one can be deformed continuously into the another without crossing itself.

Knots (and similarly links) are represented by knot diagrams, which are projections of the knot onto a plane, so that all multiple points are double

points with the strands crossing transversally. One specifies which strand crosses over. We are only concerned with smooth embeddings, in which case knots and links have finitely many crossings. Two examples are shown in Figure 1.2, the figure-eight knot and the Hopf link. It is customary to

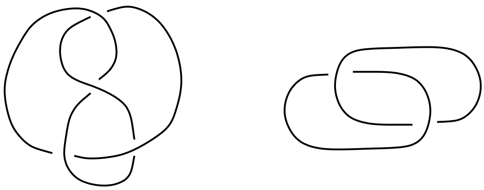


Fig. 1.2 The figure-eight knot and the Hopf link

add the point at infinity and therefore consider knots and links in the 3-dimensional sphere S^3. The theory is the same as for $\mathbb{R}^3$, but S^3 has the advantage of being a compact manifold.

In 1833, while computing the work done on a magnetic pole moving along a closed curve in the presence of a loop of current, Gauss discovered the linking number of two non-intersecting curves. He started with the Biot-Savart Law, which computes the magnetic field $\mathbf{B}$ at a given point $\mathbf{r}$ produced by an electric field of a steady current I in a thin closed wire modeled by a curve γ_1 in the space. The Biot-Savart Law states that

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi}\int_{\gamma_1} \frac{I}{\|\mathbf{r}-\mathbf{r}_1\|^3}\frac{d\mathbf{r}_1}{ds} \times (\mathbf{r}-\mathbf{r}_1)ds,$$

where μ_0 is the magnetic permeability of the vacuum and $\mathbf{r}_1(s)$ is the parametrization of γ_1. Evaluating the work of the magnetic field along a second curve γ_2 parametrized by $\mathbf{r}_2(t)$ one obtains

$$\begin{aligned}\int_{\gamma_2} \mathbf{B}\cdot\frac{d\mathbf{r}_2}{dt}dt &= \frac{\mu_0}{4\pi}\int_{\gamma_2}\int_{\gamma_1}\frac{I}{\|\mathbf{r}_2-\mathbf{r}_1\|^3}\left(\frac{d\mathbf{r}_1}{ds}\times(\mathbf{r}_2-\mathbf{r}_1)\right)\cdot\frac{d\mathbf{r}_2}{dt}dsdt \\ &= \frac{\mu_0}{4\pi}\int_{\gamma_1}\int_{\gamma_2} I\frac{\mathbf{r}_1-\mathbf{r}_2}{\|\mathbf{r}_1-\mathbf{r}_2\|^3}\cdot\left(\frac{d\mathbf{r}_1}{ds}\times\frac{d\mathbf{r}_2}{dt}\right)dtds.\end{aligned}$$

Here we used the formula $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c} = -\mathbf{b}\cdot(\mathbf{a}\times\mathbf{c})$. Setting $\mu_0 = I = 1$, Gauss produced the following formula for the linking number of γ_1 and γ_2:

$$\mathrm{lk}(\gamma_1,\gamma_2) = \frac{1}{4\pi}\int_{\gamma_1}\int_{\gamma_2}\frac{\mathbf{r_1}-\mathbf{r_2}}{\|\mathbf{r_1}-\mathbf{r_2}\|^{\mathbf{3}}}\cdot\left(\frac{d\mathbf{r_1}}{ds}\times\frac{d\mathbf{r_2}}{dt}\right)dtds,$$

where γ_1, γ_2 are the two curves parameterized by $\mathbf{r}_1$ respectively $\mathbf{r_2}$. Note that $\mathrm{lk}(\gamma_1,\gamma_2) = \mathrm{lk}(\gamma_2,\gamma_1)$, in other words, we can switch γ_1 and γ_2 and obtain the same value of the work.

As we will see below, the linking number counts the number of times one of the two curves winds around the other. The linking number will play an important role in our discussion, and is an instance of the interplay between analysis and combinatorial topology.

The value of the integral is invariant under isotopy, meaning that we are allowed to deform the curves in the 3-dimensional space as long as they don't cross each other. Much more is true, namely we have the following result.

Theorem 1.1. *If γ_1 and γ_1' are two curves such that $\gamma_1 \cup (-\gamma_1')$ bounds a surface Σ, then $lk(\gamma_1, \gamma_2) = lk(\gamma_1', \gamma_2)$.*

Proof. The proof is based on the Ampére Law, which states that the magnetic field around the boundary of a surface is proportional to the total current passing through the surface. This is a particular case of Stokes' Theorem, and our task is to show that the current passing through the surface is zero. We consider the situation where the electric current passes through γ_2.

Consider the parametrizations $\mathbf{r}_1(s) = (x(s), y(s), z_1(s))$, $\mathbf{r}_2(t) = (\tilde{x}(t), \tilde{y}(t), \tilde{z}(t))$ of γ_1 respectively γ_2 and set

$$P(x,y,z) = \int_{\gamma_2} \frac{-(\tilde{y}-y)d\tilde{z} + (\tilde{z}-z)d\tilde{y}}{((\tilde{x}-x)^2 + (\tilde{y}-y)^2 + (\tilde{z}-z)^2)^{3/2}}$$

$$Q(x,y,z) = \int_{\gamma_2} \frac{(\tilde{x}-x)d\tilde{z} - (\tilde{z}-z)d\tilde{x}}{((\tilde{x}-x)^2 + (\tilde{y}-y)^2 + (\tilde{z}-z)^2)^{3/2}}$$

$$R(x,y,z) = \int_{\gamma_2} \frac{-(\tilde{x}-x)d\tilde{y} + (\tilde{y}-y)d\tilde{x}}{((\tilde{x}-x)^2 + (\tilde{y}-y)^2 + (\tilde{z}-z)^2)^{3/2}}.$$

Set also

$$\alpha_{\gamma_2} = \frac{1}{4\pi}(Pdx + Qdy + Rdz). \tag{1.1}$$

Then

$$\mathrm{lk}(\gamma_1, \gamma_2) = \int_{\gamma_1} \alpha_{\gamma_2} = \frac{1}{4\pi}\int_{\gamma_1} Pdx + Qdy + Rdz.$$

Let $\gamma = \gamma_1' \cup (-\gamma_1)$. Stokes' theorem for γ and Σ reads

$$\int_{\gamma} \alpha_{\gamma_2} = \iint_{\Sigma} d\alpha_{\gamma_2},$$

or explicitly

$$\int_{\gamma} Pdx + Qdy + Rdz = \iint_{\Sigma} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy + \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) dydz$$
$$+ \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) dzdx.$$

We will show that the form α_{γ_2} is closed, namely that

$$\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}=\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}=\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}=0.$$

We only verify $\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}=0$, the other equalities being similar. The part of this expression containing $d\tilde{z}$ is equal to

$$\begin{aligned}&\int_{\gamma_2}-2((\tilde{x}-x)^2+(\tilde{y}-y)^2+(\tilde{z}-z)^2)^{-3/2}\\&\quad+3(\tilde{x}-x)^2((\tilde{x}-x)^2+(\tilde{y}-y)^2+(\tilde{z}-z)^2)^{-5/2}\\&\quad+3(\tilde{y}-y)^2((\tilde{x}-x)^2+(\tilde{y}-y)^2+(\tilde{z}-z)^2)^{-5/2}d\tilde{z}\\&=\int_{\gamma_2}((\tilde{x}-x)^2+(\tilde{y}-y)^2+(\tilde{z}-z)^2)^{-3/2}\\&\quad+3(\tilde{z}-z)^2((\tilde{x}-x)^2+(\tilde{y}-y)^2+(\tilde{z}-z)^2)^{-5/2}d\tilde{z}\\&=\int_{\gamma_2}\frac{\partial}{\partial\tilde{z}}((\tilde{x}-x)^2+(\tilde{y}-y)^2+(\tilde{z}-z)^2)^{-3/2}d\tilde{z}=0.\end{aligned}$$

On the other hand, only $\frac{\partial Q}{\partial x}$ has a $d\tilde{x}$ in it, so the part containing $d\tilde{x}$ in $\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}$ is equal to

$$\begin{aligned}&3\int_{\gamma_2}((x-\tilde{x})^2+(y-\tilde{y})^2+(z-\tilde{z})^2)^{-5/2}(x-\tilde{x})(z-\tilde{z})d\tilde{x}=\\&\int_{\gamma_2}\frac{\partial}{\partial\tilde{x}}\frac{z-\tilde{z}}{((x-\tilde{x})^2+(y-\tilde{y})^2+(z-\tilde{z})^2)^{3/2}}d\tilde{x}=0.\end{aligned}$$

The term involving dy' is treated similarly. The conclusion follows. $\square$

Remark 1.1. Σ can be the surface traced by γ_1 while being deformed into γ_1' but it could be any surface bounded by γ_1 and γ_1', showing that the linking number is a homological invariant. In fact $\mathrm{lk}(\gamma_1,\gamma_2)$ is the homology class of γ_1 in $H_1(S^3\backslash\gamma_2,\mathbb{Z})=\mathbb{Z}$.

As such the linking number is a link invariant for oriented links with two components (i.e. for links whose components are oriented). But more is true, namely the linking number is an integer and it can be computed combinatorially from a link diagram. To see why this is so, deform γ_1 so that it consists of several arcs connected to circles that link γ_2, such as in Figure 1.3. As the arcs are traveled back and forth, it suffices to replace γ_1 with the union of the circles. We can compute the linking number separately for each circle, and then add up.

Next, we can deform γ_2 so that everything besides the arc that links with the circle γ_1 is pushed towards infinity, and that arc and the circle

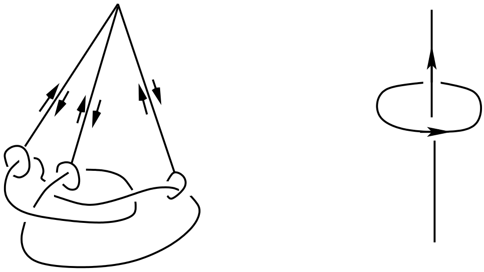


Fig. 1.3 The computation of the linking number

look like the z-axis and the circle $x^2 + y^2 = 1$ in $\mathbb{R}^3$. Parametrizing γ_2 by $\mathbf{r}_2 = (0, 0, t)$, there are two possibilities for the parametrization of γ_1:

$$\mathbf{r}_1 = (\cos s, \sin s, 0) \text{ or } \mathbf{r}_2 = (\cos(-s), \sin(-s), 0).$$

The first case corresponds to the situation from Figure 1.3, and we compute

$$\begin{aligned} \mathrm{lk}(\gamma_1, \gamma_2) &= \frac{1}{4\pi}\int_{-\infty}^{\infty}\int_0^{2\pi} \frac{(\cos s, \sin s, -t)}{(1+t^2)^{3/2}} \cdot [(-\sin s, \cos s, 0) \times (0,0,1)]dsdt \\ &= \frac{1}{4\pi}\int_{-\infty}^{\infty}\int_0^{2\pi} \frac{1}{(1+t^2)^{3/2}}dsdt = \frac{1}{2}\,\frac{t}{\sqrt{1+t^2}}\Big|_{-\infty}^{\infty} = 1. \end{aligned}$$

In the other case the value of the integral is -1.

This allows a combinatorial computation of the linking number from a link diagram. Define the sign of a crossing using Figure 1.4, with the crossing on the left being positive and the one on the right negative. In a link diagram, each crossing is of one of these two types. Consider only the crossings where γ_1 crosses over γ_2, and let n_1 and n_2 be the number of positive, respectively negative crossings. Then

$$\mathrm{lk}(\gamma_1, \gamma_2) = n_1 - n_2.$$

If we perform the computation using both the over and the under crossings, we obtain twice the linking number.

Remark 1.2. The above computation shows that if D is an oriented disk that intersects the closed curve γ at one point, and if ∂D is oriented so that D is on the left, then $\mathrm{lk}(\partial D, \gamma) = \pm 1$. The sign is $+1$ if the 3-dimensional frame obtained by adjoining to an orientation frame of D the tangent vector of γ is positively oriented and -1 if this frame is negatively oriented.

Example 1.2. Depending on the orientation of the link components, the linking number of the Hopf link (Figure 1.2) is either 1 or -1.

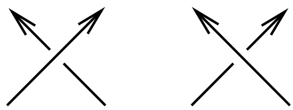


Fig. 1.4 A positive and a negative crossing

If we are given just one knotted oriented curve, γ, and consider a knot diagram of it, then we can use the same rules for the crossings in this diagram. The difference between the number of positive crossings and the number of negative crossings is called the **writhe** of the knot diagram. The writhe is not a knot invariant, since it changes if we introduce twists (kinks) like those shown in Figure 1.5.

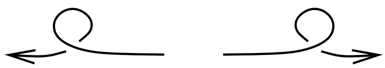


Fig. 1.5 Positive and negative twists

If we are given a link with two components γ_1 and γ_2, then we can compute $\mathrm{lk}(\gamma_1, \gamma_2)$ from a link diagram, by taking the difference between the total number of positive crossings and that of negative crossings, then subtracting the sum of the writhes of the projections of γ_1 and γ_2, and then dividing the answer by 2.

1.2.2 *The Jones polynomial*

In 1984, Vaughan F.R. Jones discovered a polynomial invariant of knots, which can be computed recursively [Jones (1985)]. For a knot K, the Jones polynomial is a one-variable polynomial, $V_K(t)$, that is equal to 1 for the unknot, and can be computed for any other knot by orienting it and then applying repeatedly the relation

$$t^{-1}V_{K_+}(t) - tV_{K_-}(t) = (t^{1/2} - t^{-1/2})V_{K_0}(t). \tag{1.2}$$

The relation (1.2) is called a **skein relation**. In this expression, K_+, K_-, K_0 denote three oriented knots or links which have the same diagram except for a crossing, and that crossing is positive for K_+, negative for K_-, while for K_0 the crossing is erased and the strands are connected so that the orientations agree.

Definition 1.1. A skein relation for a knot and link invariant is a linear relation among the invariants of the knots and links obtained by modifying one crossing according to some rules.

In the case of the Jones polynomial, the crossing is modified by changing it from an undercrossing to an overcrossing or vice-versa, and by cutting it open and joining the ends so that the crossing disappears.

Example 1.3. The computation of the Jones polynomial for the right-handed trefoil knot is shown in Figure 1.6.

One obtains

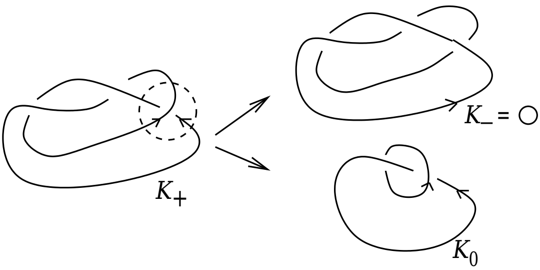
$$\begin{aligned} V_{\text{trefoil}}(t) &= t^2 + (t^{3/2} - t^{1/2})V_{\text{Hopf}}(t) \\ &= t^2 + (t^{3/2} - t^{1/2})[t^2(-t^{1/2} - t^{-1/2}) + (t^{3/2} - t^{1/2})] \\ &= t + t^3 - t^4. \end{aligned}$$

Jones discovered his polynomial while using combinatorial methods for understanding how an algebra of quantum observables lies inside another. No intrinsic topological definition of the Jones polynomial is known at the date of publication of this book. That the Jones polynomial is a topological invariant, namely that it is independent of which projection of the knot is considered, follows by checking its invariance under **Reidemeister moves**. These are the following:

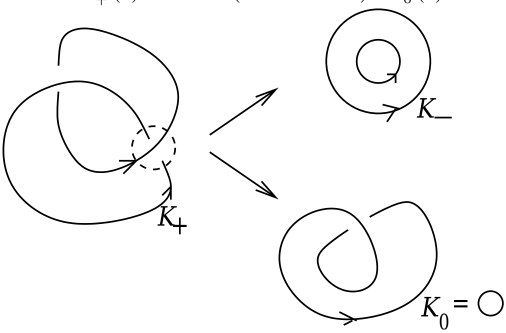
I. the elimination or addition of a twist,
II. the separation or overlapping of two strands,
III. the passing of a strand over/under/between two crossing strands.

Figure 1.7 depicts one Reidemeister move of each type. Reidemeister's theorem [Reidemeister (1926)] states that any two diagrams of the same knot and link can be changed into one another by a finite sequence of such moves.

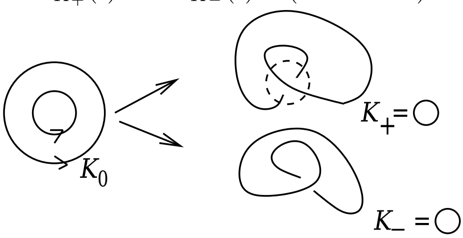
Any quantity that is associated to knots and links and is invariant under the Reidemeister moves is a knot and link invariant. In particular, we could have defined the linking number by counting the positive and negative crossings, as in §1.2.1, and then check invariance under Reidemester moves in order to prove that it is a topological invariant.



$$V_{K_+}(t) = t^2 + (t^{3/2} - t^{1/2})V_{K_0}(t)$$



$$V_{K_+}(t) = t^2 V_{K_-}(t) + (t^{3/2} - t^{1/2})$$



$$V_{K_0}(t) = -t^{1/2} - t^{-1/2}.$$

Fig. 1.6 Computation of the Jones polynomial of the trefoil knot

1.2.3 *Computing the linking number from skein relations*

Returning to the linking number, by analogy with the Jones polynomial we can also compute it from skein relations in a link diagram. For this we introduce the notion of framed oriented links.

Definition 1.2. A **framing** of a smooth knot $\gamma : S^1 \to \mathbb{R}^3$ is a choice of a smooth vector valued function $f : S^1 \to \mathbb{R}^3\backslash\{0\}$ such that for each $\tau \in S^1$, $f(\tau)$ is orthogonal to the tangent vector $\gamma'(\tau)$. A knot endowed with a framing is called a **framed knot**.

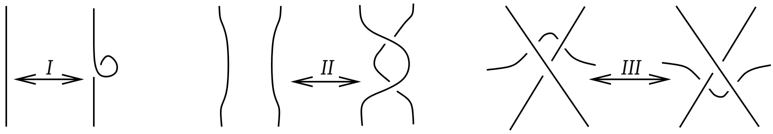


Fig. 1.7 The Reidemeister moves

In a more intuitive language, a framed knot is an embedded ribbon; just view the vector $f(\tau)$ as a little arrow attached to the point $\gamma(\tau)$, $\tau \in S^1$, and then the ribbon is determined by these vectors. This is not entirely correct, since $-f$ defines the same ribbon, but the two can be changed one into the other by spinning the framing 180° about the knot. Note that physical knots made out of rope are framed by nature, the rope plays the same role as the ribbon in keeping track of the twistings of the knot around itself.

Definition 1.3. A **framed link** is a link whose components are framed knots.

Given a framed link and a plane, the link can always be deformed continuously without crossing itself so that the framing becomes parallel to the plane. In this case if we are given a link diagram we can recover the framing from the diagram, by taking a regular neighborhood of the projection in the plane. We say that the link defined by the diagram has the **blackboard framing**.

Reidemeister's Theorem can be adapted to show that two diagrams represent the same framed link if they can be transformed into one another by Reidemeister II and III moves. However, Reidemeister I move changes the framing.

We agree that a link diagram consisting of only unknotted disjoint circles is associated the constant polynomial equal to 1.

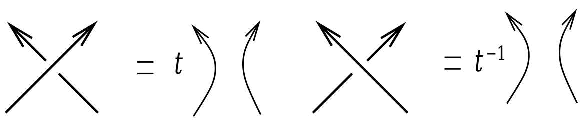


Fig. 1.8 Skein relations for the linking number

Given a link diagram of an oriented link L with blackboard framing, apply the **skein relations** from Figure 1.8 until the diagram consists only

of unknotted disjoint circles, then replace these circles by the constant polynomial equal to 1. The result is the polynomial $t^{\mathrm{lk}(L)}$ for some integer $\mathrm{lk}(L)$. There is an integral formula for computing the exponent $\mathrm{lk}(L)$.

To explicate it, define the link $L_{\rightarrow}$ to be the parallel copy of L in the direction of the framing. Alternatively, we can think that L and $L_{\rightarrow}$ contain respectively the boundary components of the annuli that comprise the framed link. Recall the form α_γ defined in (1.1). Define the 1-form $\alpha_L = \sum \alpha_\gamma$, where γ ranges over the components of L. Then

$$\mathrm{lk}(L) = \int_{L_{\rightarrow}} \alpha_L = \sum \int_{\gamma_{\rightarrow}} \alpha_L,$$

the sum being taken over all components of $L_{\rightarrow}$. Stokes' Theorem implies that $\mathrm{lk}(L)$ is a framed oriented link invariant.

If L consists of two components, γ_1 and γ_2, then

$$t^{\mathrm{lk}(L)}/(t^{\mathrm{lk}(\gamma_1)}t^{\mathrm{lk}(\gamma_2)}) = t^{2\mathrm{lk}(\gamma_1,\gamma_2)}.$$

For this reason we will call the skein relations from Figure 1.8 the skein relations of the linking number.

1.3 Witten's Chern-Simons theory

A few years after Jones made his discovery, Edward Witten explained the Jones polynomial using quantum field theory [Witten (1989)]. This was done in order to give an intrinsic definition to the Jones polynomial independent of knot diagrams. Let us briefly explain Witten's work, so that the reader will understand some ideas that led to the writing of this book. We will return with a few more details in Chapter 9.

Let G be a compact Lie group with Lie algebra $\mathfrak{g}$. This is the gauge group of the theory. Consider the fields with symmetry group G on a 3-dimensional manifold M without boundary. The presence of such a field is determined by its action on the phase of a particle moving through M. The phase is described by a vector, which is rotated by an element of G.

Each field has a potential, which is a $\mathfrak{g}$-connection A in the trivial principal bundle $M \times G$. As such, A is defined by a $\mathfrak{g}$-valued 1-form, denoted also by A. If we integrate the connection along a loop, we obtain an element of G, by which the phase of the particle is rotated when it travels along that loop. This element of G is called the **holonomy** of the connection along the loop. In Witten's model, the classical observables are the **traces of the holonomies** of such connections in a certain representation of the

Lie group. So, from the point of view of classical physics, the measurable quantities are the functions of fields

$$A \mapsto \operatorname{tr}_V \operatorname{hol}_\gamma(A)$$

where γ is a loop and V is a representation of G. These are called **Wilson lines**.

This theory has a Lagrangian, the **Chern-Simons functional**

$$L(A) = \frac{1}{4\pi}\int_M tr(A \wedge dA + \frac{2}{3}A \wedge A \wedge A),$$

which was introduced by S.S. Chern and J. Simons in [Chern and Simons (1974)].

Witten outlined a way of producing a quantum theory with this Lagrangian, using Feynmann integrals. Planck's constant is chosen to be the reciprocal of an integer: $h = \frac{1}{N}$. To the classical observable that is the Wilson line defined by the curve γ and the representation V one associates the Feynmann integral

$$\int e^{\frac{i}{h}L(A)} \operatorname{tr}_V \operatorname{hol}_\gamma(A)\mathcal{D}A. \tag{1.3}$$

This integral is taken over the infinite dimensional space of all connections. As such, it is not well defined mathematically. It should be thought of as an average of the quantities $\operatorname{tr}_V \operatorname{hol}_\gamma(A)$ over all connections A, where the average is weighted by $e^{i/hhL(A)}$.

Witten claimed that for $G = SU(2)$, the manifold M equal to S^3, γ a knot K, and V the standard 2-dimensional representation of $SU(2)$, the value of the integral (1.3) equals the Jones polynomial evaluated at a root of unity. That is

$$\int e^{iNL(A)} \operatorname{tr}_V \operatorname{hol}_\gamma(A)\mathcal{D}A = V_K\left(e^{\frac{\pi i}{N}}\right).$$

While this fact cannot be established rigorously, certain properties of Feynmann integrals suggest that this is indeed so. These properties allow the localization of the computation, by cutting the manifold M into pieces, computing the integral on each piece, and then combining the results according to certain rules. A crossing can then be placed inside a ball. The Feynmann integral on a ball with two strands inside would be a 2-dimensional vector. Hence the three vectors associated to the ball of the crossing in K_+, K_- respectively K_0 are linearly dependent. The linear dependence is the skein relation of the Jones polynomial.

Of interest to us, in this book, is the case

$$G = U(1) = \{z \in \mathbb{C} \mid |z| = 1\}.$$

This case is known as the **abelian Chern-Simons theory**, and was first studied by Albert Schwarz in [Schwarz (1978)]. For similar reasons this case would be related to the linking number. Here, the vectors associated to a crossing are 1-dimensional, the linear dependence of two such vectors giving rise to the skein relations from Figure 1.9.

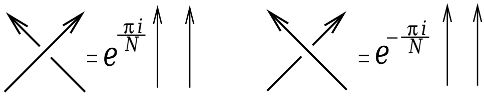


Fig. 1.9 Skein relations for $U(1)$-knot invariants

While Witten's constructs lacked rigorous foundation, they gave rise to rich mathematics. First, everything depends only on the topology of the manifold M and of the embedded loop γ, so we are in the presence of a **topological quantum field theory** (TQFT). Such theories have been formalized by M.F. Atiyah [Atiyah (1988)], [Atiyah (1990)]. In particular,

$$Z(M) = \int e^{\frac{i}{\hbar}L(A)} \mathcal{D}A$$

is a topological invariant of the manifold M.

Then, the invariance of Witten's integrals under isotopies of the curve γ means invariance under Reidemeister moves. The third of these moves relates to the Yang-Baxter equation in statistical dynamics, and hence to quantum groups.

Both the Chern-Simons Lagrangian, and the Wilson line, are invariant under the changes of coordinates of the field. These changes of coordinates are called gauge transformations. They are defined by smooth functions $g : M \to G$, and change the potential by

$$A \mapsto g^{-1}Ag + g^{-1}dg.$$

Witten's integrals are therefore taken over equivalence classes of connections.

The possibility of computing Witten's integrals by decomposing the manifold into pieces gives rise to an extension of these integrals to manifolds with boundary. To a manifold with boundary corresponds a vector in a finite dimensional vector space that is associated to the boundary surface. These vector spaces arise from the quantization of the space of $\mathfrak{g}$-connections

on the boundary surface. Using the symmetries of the system and symplectic reduction, one reduces the quantization problem to the quantization of the space of **flat $\mathfrak{g}$-connections (those with curvature zero) on a surface, modulo gauge transformations**. This moduli space is now a finite dimensional space, which is endowed with the structure of a finite dimensional algebraic variety. The smooth part of this variety has all the good properties of the phase space of a classical mechanical system. The problem now becomes quantum mechanical, and the standard quantization procedure is to replace the moduli space of connections by a Hilbert space, and the Wilson lines of curves on the surface by linear operators on this Hilbert space.

In the case $G = U(1)$, if we endow the surface that is the boundary of the manifold with a complex structure, then the moduli space of flat connections becomes a complex torus, which turns out to be the Jacobian variety of the surface. The Hilbert space of the quantum mechanical system is the space of Riemann's theta functions, and this is how theta functions enter the picture. Witten's considerations therefore show that there is a close relationship between theta functions and the linking number of knots. The aim of this book is to put this relationship in a perspective different from Witten's, using representation theory, and as such, to avoid the heuristical considerations of quantum field theory.

Throughout the book, whenever we refer to Chern-Simons theory, we mean the constructs of Edward Witten mentioned above and the mathematics they gave rise to.