Máster en Física Teórica
Facultad de Ciencias
Universidad Autónoma de Madrid

Master Thesis

Madrid 2016

Tachyonic instabilities in Yang-Mills theories and number theory

Fernando Chamizo Lorente

Advisor: **Antonio González-Arroyo España**

I said [to S.S. Chern] I found it amazing that gauge theory are exactly connections on fibre bundles, which the mathematicians developed without reference to the physical world. I added "this is both thrilling and puzzling, since you mathematicians dreamed up these concepts out of nowhere:" He immediately protested: "No, no. These concepts were not dreamed up. They were natural and real." C.N. Yang in [80].

Contents

Preface

Each researcher is under continued training, learning from books, papers and colleagues. In principle, any of them should be avid for being taught, but I have learned the hard way why official education is addressed to young people. This work is the culmination of a exhausting year under not very supportive personal circumstances.

Various things, nevertheless, have been on my side. One of them is that by sheer chance, I have worked on a problem connecting theoretical physics and my previous research topic. I thank my advisor, Antonio González-Arroyo, for his enthusiasm sharing and proposing the problem, and for his patience and work.

Except for the section devoted to the number theoretical approach, the rest of the main contents can be found in the specialized literature. In particular, the $2+1$ Yang-Mills model studied here comes from [29] and [30]. If all goes as planned, we will compose a research paper [9] with the new approach. In fact the initial plan was to proceed in the other direction but my poor performance and slowness prevented it. I want to end this short preface remembering Javier Cilleruelo, who unexpectedly passed away recently. I am sure that he would have enjoyed to know that the kind of problems he loved may appear in theoretical physics.

Chapter 1

Basic ideas about gauge theories

1.1 Elementary examples and concepts

A motivating example

Let us start with a very illustrative and simple example. It is taken from [3] (in an abbreviated form) and, as far as we know, it is not a historical example, but looking into retrospective it has some connections with the ideas introduced in 1926 by V. Fock [42] shortly after E. Schrödinger stated his equation. We are going to use the kind of *first quantization* arguments that a pioneer of quantum mechanics could have employed at that time.

After the classical contribution of J.C. Maxwell and H. Lorentz, it was known that the *Lorentz force* $q(\vec{E}+\vec{v}\times\vec{B})$ on a charge q derives from a classical Hamiltonian

$$\mathcal{H} = \frac{1}{2m}\big(\vec{p} - q\vec{A}\big)^2 + q\varphi \tag{1.1}$$

where $(A_\alpha) = (\varphi, -\vec{A})$ is what we call today the 4-potential. We have

$$\vec{E} = -\nabla\varphi - \frac{\partial\vec{A}}{\partial t} \qquad \text{and} \qquad \vec{B} = \nabla\times\vec{A}. \tag{1.2}$$

This equations, that embody two of the *Maxwell equations*, the Lorentz force and consequently the dynamics of the system, are invariant by the *gauge transformation*

$$\varphi \longmapsto \varphi + \frac{\partial\chi}{\partial t}, \qquad \vec{A} \longmapsto \vec{A} - \nabla\chi. \tag{1.3}$$

On the other hand, the (first) quantization of this electromagnetic problem via *Schrödinger's equation* is (in natural units $\hbar = c = 1$)

$$\Big(\frac{1}{2m}\big(-i\nabla - q\vec{A}\big)^2 + q\varphi\Big)\Psi = i\frac{\partial\Psi}{\partial t}. \tag{1.4}$$

There is something strange in this equation: It seems that the transformations (1.3) change it substantially and then the choice of different gauges could give, in principle, different physical conclusions. But χ in (1.3) is a mathematical artifact related to the non uniqueness of the solution of the partial differential equations $\nabla \cdot \vec{F} = f_0$ and $\nabla \times \vec{F} = F_0$, and there is not a clear rule to privilege a particular solution. At least in the classic setting the value of the 4-potential at a point has not physical significance. At a fixed point, we can only measure the electric and magnetic fields $\vec{E}$ and $\vec{B}$.

It turns out that the gauge change (1.3) that apparently modifies completely (1.4) does not act dramatically on the solutions. It can be checked that if Ψ is a solution of (1.4) for certain $(\varphi, -\vec{A})$ then $e^{-iq\chi}\Psi$ is a solution after the gauge change (1.3) (shortly we shall see the calculation in detail in another example). With this information at hand, one can argue that the "probability density" $|\Psi|^2$ is invariant under phase changes and take it as an explanation to save the gauge invariance coming from the well-settled Maxwell equations. But this explanation has a flaw, if we allow a possibly time dependent phase in the wave function, then the probability interpretation and its conservation, is in danger. Note that $\Psi^*\nabla\Psi - \Psi\nabla\Psi^*$ is no longer a current, as in the case of a free particle. The important point to be noted here is that (1.4) becomes the free particle Schrödinger's equation if

$$\nabla - iq\vec{A} \longmapsto \nabla \qquad \text{and} \qquad \frac{\partial}{\partial t} + iq\chi \longmapsto \frac{\partial}{\partial t}. \tag{1.5}$$

This sounds relativistic and it is noteworthy in our context in which relativity was not explicitly considered. The important point is that it gives a big clue about the right conserved current and, in general, an answer about why different gauges lead to the same physics: The operator

$$D = \Big(\frac{\partial}{\partial t}. + iq\varphi, \nabla - iq\vec{A}\Big) \tag{1.6}$$

is in some sense gauge invariant, meaning that if D' is the operator in other gauge, and $\Psi' = e^{-iq\chi}\Psi$ is the solution of (1.4) in that gauge, then

$$D'\Psi' = e^{-iq\chi}D\Psi. \tag{1.7}$$

Let us insist on the same point reviewing with care the computations in a relativistic example to avoid any paradox coming from the combination of Newton's dynamics and electromagnetism. Imagine that, as Schrödinger tried in first place [26], we want to study the quantum relativistic corrections for a charged particle. The natural quantization for the uncharged free particle in natural units ($c = \hbar = 1$) is

$$E^2 + \vec{p}^2 = m^2, \qquad E \leftrightarrow i\frac{\partial}{\partial t}, \quad \vec{p} \leftrightarrow -i\nabla, \tag{1.8}$$

that leads to the Klein-Gordon equation

$$\partial_\alpha \partial^\alpha \Psi + m^2 \Psi = 0. \tag{1.9}$$

If we now switch the electromagnetic field on, one should add the corresponding potential energy to the 4-moment, i.e. one should replace $i\partial_\alpha$ by $i\partial_\alpha + qA_\alpha$. Let us write $q = -e$ having in mind the electron. Consequently, (1.9) becomes

$$\big(\partial_\alpha + ieA_\alpha\big)\big(\partial^\alpha + ieA^\alpha\big)\Psi + m^2\Psi = 0. \tag{1.10}$$

Again, if Ψ solves (1.10) then $e^{-ie\chi(x)}\Psi$ solves (1.10) after the gauge change (1.3). The key point is

$$\begin{cases} \big(\partial^\alpha + ie(A^\alpha + \partial^\alpha\chi)\big)\big(e^{-ie\chi(x)}\Psi\big) = e^{-ie\chi(x)}\big(\partial_\alpha + ieA^\alpha\big)\Psi, \\ \big(\partial_\alpha + ie(A_\alpha + \partial_\alpha\chi)\big)\big(e^{-ie\chi(x)}\Psi\big) = e^{-ie\chi(x)}\big(\partial_\alpha + ieA_\alpha\big)\Psi. \end{cases} \tag{1.11}$$

There is nothing deep in these relations, we simply compensate the extra term in the derivative of a product using $\big(\partial_\alpha - \partial_\alpha f\big)e^f = 0$, the defining property of the exponential.

We can express the situation in a very succinct way introducing the *covariant derivative* D and the *gauge transformation* G

$$D_\mu = \partial_\mu + ieA_\mu \qquad \text{and} \qquad G\Psi = e^{-ie\chi(x)}\Psi. \tag{1.12}$$

Then our observation is that under $\Psi \longmapsto G\Psi$, we have

$$D_\mu\Psi \longmapsto GD_\mu\Psi \qquad \text{and} \qquad A_\mu\Psi \longmapsto GA_\mu G^{-1} + ie^{-1}(\partial_\mu G)G^{-1}. \tag{1.13}$$

Of course, the second equation is just verbosity meaning simply $A_\mu \mapsto A_\mu + \partial_\mu\chi$.

Extending these ideas to the right context of spin 1/2 particles (to include the electron) leads to write the QED Lagrangian as

$$\mathcal{L} = \bar\Psi(i\not{D} - m)\Psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \qquad \text{with} \quad \not{D} = \gamma^\mu D_\mu. \tag{1.14}$$

It is clearly invariant by (1.13). Note that the interaction Lagrangian $-j^\mu A_\mu = -e\bar\Psi\gamma^\mu\Psi A_\mu$ is obtained by *minimal coupling* changing usual derivatives ∂_μ by covariant derivatives D_μ. In other words, from the *Dirac equation* for the free particle[1].

[1] In the case of the electromagnetic field, this is not so spectacular. In the famous paper [16] in which P.A.M. Dirac introduces his equation, we can read "we adopt the usual procedure of substituting $p_0 + e/cA_0$ for p_0 and $\mathbf{p} + e/c\mathbf{A}$ for $\mathbf{p}$ in the Hamiltonian for no field."

A non-Abelian example

Let us consider now an example with less physical significance but closer to the ideas appearing in the Standard Model.

Say that we have N real scalar fields $\phi_1, \phi_2, \dots \phi_N$ behaving as harmonic oscillators with equal mass and no interaction. The Lagrangian is

$$\mathcal{L} = \frac{1}{2}\sum_{k=1}^{N} \partial_\mu \phi_k \partial^\mu \phi_k - \frac{1}{2} m^2 \sum_{k=1}^{N} \phi_k^2. \tag{1.15}$$

We can define artificially a column vector field Φ in $\mathbb{R}^N$ having ϕ_k as its k-th coordinate, and write

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \Phi)^\dagger (\partial^\mu \Phi) - \frac{1}{2} m^2 \Phi^\dagger \Phi \tag{1.16}$$

(of course, the dagger can be changed by transposition).

There is no a priori physical reason to consider the different fields as components of the same object. In the same way, we could say that neutrinos and electrons are different animals and coupling both together in the Standard Model Lagrangian [74] is, in principle, an economic usage of the mathematical notation like in (1.16).

The group $SO(N)$ acts naturally on Φ preserving (1.16). In fact, $O(N)$ also leaves it invariant, but by technical reasons, we only consider the connected component containing the identity. Each transformation of $SO(N)$ mixes the real fields ϕ_k together in the same way as $SU(2)$ mixes neutrinos and electrons. Physics shows up when we assume a kind of *gauge principle*, claiming a minimal coupling rule. Shortly we shall see it in a broader and more formal context. Now we are going to tackle the problem through our example.

In the electrodynamic example, the formalism allows us to interpret the interacting field as a term to be added to the derivative in such a way that the result behaves well under the symmetry transformations, even if they vary from a point to another. Then we consider as in (1.12)

$$D_\mu = \partial_\mu + igA_\mu \qquad \text{and} \qquad G\Psi = S(x)\Psi. \tag{1.17}$$

where $S(x) \in SO(N)$ for each x. Now A_μ, the *gauge field* is at each point an $N \times N$ matrix, and g is just a constant introduced to recall the physical situations ($g = e$ in the previous case).

If we wish the covariant derivative D_μ to be really covariant, $D_\mu S(x) = S(x) D_\mu$, then A_μ must transform in the particular way stated in (1.13) with $e = g$. Both formulas in (1.13) become equivalent. To keep a complete analogy one would like to write $S(x)$ as an exponential $G\Psi = e^{-igX(x)}\Psi$ where $X(x)$ is a purely imaginary anti-symmetric matrix to assure $e^{-igX(x)} \in SO(N)$. We can identify then the exponents with elements of the Lie algebra $\mathfrak{so}(N)$. This simply reflects that derivatives in a Lie group give rise to elements in the Lie algebra. Expanding this comment, since ∂_μ and A_μ stand with the same role in (1.17), it is natural to impose $A_\mu \in \mathfrak{so}(N)$. Hence we could express each A_μ as a linear combination of the $N(N-1)/2$ elements of a fixed basis of $\mathfrak{so}(N)$. On the other hand, μ varies in $\{0,1,2,3\}$, then we can think the gauge field as a superposition of $N(N-1)/2$ vector fields (with matrix coefficients). In the Standard Model, the strong force corresponds to the group $SU(3)$ (there are three colors), then there are $\dim \mathfrak{su}(3) = 3^2-1$ basic gauge (gluon) fields represented by the Gell-Mann matrices.

Under the "spell" [69] of the gauge principle, we can guess that in our example the Lagrangian including the interaction term is

$$\mathcal{L} = \frac{1}{2}(D_\mu\Phi)^\dagger(D^\mu\Phi) - \frac{1}{2}m^2\Phi^\dagger\Phi. \tag{1.18}$$

Note that, unwrapping the notation, it gives an interaction Lagrangian that is not trivial to guess

$$\mathcal{L}_{\text{int}} = \frac{1}{2}ig(A_\mu\Phi)^\dagger\partial^\mu\Phi + \frac{1}{2}ig(\partial^\mu\Phi)^\dagger A_\mu\Phi - \frac{1}{2}g^2(A_\mu\Phi)^\dagger A^\mu\Phi. \tag{1.19}$$

If we compare it to the electromagnetic example, this is not the whole story. In (1.14) we had the term $F_{\mu\nu}F^{\mu\nu}$ giving the Maxwell equations when taking variations [27]. Its proxy in the non-Abelian setting is the Yang-Mills term.

Some historical geometric aspects and Yang-Mills theories

The inaugural lecture "On the hypothesis which lie at the foundations of geometry" given by B. Riemann in 1854 is considered a landmark in mathematics (see [62] for an English translation with detailed explanations). It is known that C.F. Gauss praised it greatly[2]. But it was not published until the year in which Riemann passed away and the first impression for a reader is that of a vague outreach paper with very few formulas. It only reflects the hope of Riemann to be understood by the most of the audience.

A main point in that talk is that one can define intrinsically geometric objects by local metric properties, without any reference to an outer space. It took many years and many authors to develop this idea, that was extremely important in general

[2]He forced the topic breaking the tradition of admitting the candidate first choice that was Riemann's Habilitationsschrift, a masterpiece on Fourier series.

relativity. One of the pioneer contributors to this new view of geometry was T. Levi-Civita who defined an *absolute derivative* [46] introducing a method to displace a vector to a nearby point, this was a way to "connect" points and the origin of the term *connection*. Without entering into the mathematical definitions, the idea is very simple: If there is not a privileged global orthonormal frame, to study the variation (the derivative) of a vector field V, we have to take into account the variation of the coordinates plus the variation of the reference frame, this gives *covariant derivatives* or *connections* (in a strict sense)

$$D_i V = \big(\partial_i V^k + \Gamma_{ij}^k V^j\big)\partial_k. \tag{1.20}$$

The functions Γ_{ij}^k must satisfy certain consistency conditions to avoid contradictions when changing coordinates. It turns out that there is only a possible choice of the Γ_{ij}^k if one wishes to keep some compatibility with the metric structure [62]. This is the so-called *Levi-Civita connection* in which Γ_{ij}^k are the Christoffel symbols.

It is apparent the similarity between (1.20) and the covariant derivative in the context of gauge fields. This similarity becomes identity when the concept of *vector bundle* or, more in general, that of *fiber bundle* is taken from mathematics. Surprisingly, it seems that, after the breakthrough in the physical gauge theory by C.N. Yang and R. Mills in 1954 [78], it took many years to note this coincidence. The following opinion was expressed by Yang in a recent colloquium [77]

> It came as a great shock [...] when it became clear in the 1970s that the mathematics of gauge theory, both Abelian and non-Abelian, is *exactly* the same as that of fiber bundle theory. [...] it served to *bring back* the close relationship between the two disciplines [mathematics and physics] which had been interrupted through the increasingly abstract nature of mathematics since the middle of the 20th century.

Without entering into details, we are going to give a formulation of what can be called a *gauge principle* with a mathematical flavor (but not very technical). Let G be a generic element of a matrix Lie group included in $GL(N)$ and consider the covariant derivative

$$D_\mu \Psi = \partial_\mu \Psi + igA_\mu \Psi \qquad \text{with } A_\mu \text{ in the Lie algebra.} \tag{1.21}$$

Assume for simplicity that $\Psi : \mathcal{U} \longrightarrow \mathbb{R}^N$ with $\mathcal{U} \subset \mathbb{R}^n$ (more formally, Ψ should be a *section* of the bundle). Under

$$\Psi \longmapsto G\Psi \qquad \text{and} \qquad A_\mu \longmapsto GA_\mu G^{-1} + ig^{-1}(\partial_\mu G)G^{-1} \tag{1.22}$$

we have that $D_\mu\Psi \mapsto GD_\mu\Psi$, i.e. the derivative is actually covariant. Here g is a *coupling constant* that is separated from A_μ for convenience in the physical interpretation. In our first examples was the charge of the electron.

Let us write G_0 to indicate a generic element of the Lie group constant in $\mathcal{U}$. Then we can deduce that for any Lagrangian $\mathcal{L} = \mathcal{L}(D_\mu\Psi, \Psi)$,

(1.23)
$$\mathcal{L}(\partial_\mu\Psi, \Psi) \text{ invariant under } \Psi \mapsto G_0\Psi \quad \Rightarrow \quad \mathcal{L}(D_\mu\Psi, \Psi) \text{ invariant under (1.22).}$$

As in the examples above, (1.23) can be used to create involved Lagrangians once one guesses the *gauge group*. This idea has been specially successfully in the creation and development of the Standard Model. We can rephrase (1.23) in physical terms saying that the invariance of the Lagrangian with respect to local transformations determines the interaction of Ψ with the gauge field A_μ. But we still need a term, like $F_{\mu\nu}F^{\mu\nu}$ in (1.14), giving the Lagrangian of the field itself.

This new term must depend only on A_μ and must be gauge invariant. In Riemannian geometry an important construction associated to the Levi-Civita connection and already appearing in Riemann's inaugural talk is the *curvature tensor*. It corresponds to the difference between cross partial covariant derivatives. Define $F_{\mu\nu}$ such that for any Ψ

$$[D_\mu, D_\nu]\Psi = igF_{\mu\nu}\Psi. \tag{1.24}$$

Note the analogy with Faraday's tensor. It can be proved that it is well defined and behaves as a covariant tensor[3] in the indexes μ and ν. On the other hand, under gauge transformations on A_μ, as in the second part of (1.22), it changes as $F_{\mu\nu} \mapsto GF_{\mu\nu}G^{-1}$ (this is simpler than it seems, just the covariance of the derivatives). Then the natural "scalar" $F_{\mu\nu}F^{\mu\nu}$, actually a matrix, obeys the same rule and we can do it invariant under the gauge transformations taking traces. Summing up, a plausible term to add in the Lagrangian is the *Yang-Mills Lagrangian*

$$\mathcal{L}_{\text{YM}} = -\frac{1}{2}\text{Tr}\big(F_{\mu\nu}F^{\mu\nu}\big). \tag{1.25}$$

The Euler-Lagrange equations corresponding to (1.25) are the *Yang-Mills equations* [51]

$$\partial_\mu F_{\mu\nu} + ig[A_\mu, F_{\mu\nu}] = 0, \tag{1.26}$$

that in the case of the electromagnetic field are the Maxwell equations.

Ironically, this cumbersome geometrization of gauge theories in which the potential corresponds to the connection and the field strength to the curvature [31] that has been crucial in the modern understanding of the Standard Model, is very close

[3] The usual introduction of the curvature tensor in Riemannian geometry is as a tensor field covariant in three indexes and contravariant in one. To recover the analogy one has to note that in our setting $F_{\mu\nu}$ is a matrix, a tensor of type $(1,1)$, at each point.

to the initial introduction and development in the period 1918–1929 of the concept of gauge done by H. Weyl. His aim was the unification of the electromagnetism and gravitation (general relativity), where the curvature plays an important role in the theory. In his works the term "gauge", introduced in 1918 by himself (in German), is more natural than nowadays because it corresponded to a change of scale in the metric [53]. The modern language of differential forms initiated by É. Cartan is specially useful to present these ideas in a compact form. We have avoided it because it requires a stronger background.

A final comment is that the electro-weak sector in the Standard Model presents an important variation with respect the scheme explained here. It is a Yang-Mills theory with group $U(1) \times SU(2)$ giving $1 + \dim \mathfrak{su}(2) = 4$ fields, associated to the photon and the bosons $W^\pm$ and Z. The problem is that the massive nature of the particles $W^\pm$ and Z (related to the short range of the weak interaction) does not fit the pure Yang-Mills scheme. A new term was added to the Lagrangian to solve this problem breaking the symmetry imposed by the gauge group (see [24, §9.3] for a simple and convincing mathematical explanation). It is associated to the *Higgs particle* that was recently detected at the LHC.

1.2 Quantizing Yang-Mills theory with a lattice

The path integral

The path integral formalism is one of the highlights of quantum mechanics. It can be traced back to a paper by Dirac [17] that, as its title suggests, in some way recovers the Lagrangian for a quantum mechanic dominated by the Hamiltonian, the main mechanical ingredient in the Schrödinger's equation. This paper was unnoticed during many years (it was published in a less known journal, fortunately J. Schwinger includes it in his selection [58]). The idea was retaken and fully developed by R. Feynman. Probably the prestige and social projection of Feynman, who wrote the wonderful popular science book on this idea [22], have eclipsed Dirac's contribution even today, although Feynman referred to it in his Nobel lecture and in other occasions.

There are many specific examples in [23], we simply mention the general idea following mainly [50].

Consider the classical time independent Hamiltonian in one spatial dimension

$$H = \frac{p^2}{2m} + V \qquad \text{with} \quad V = V(x), \quad x \in \mathbb{R}. \tag{1.27}$$

Schrödinger's equation tells us that the probability amplitude to go from (x_0, t_0) to

(x_1, t_1) is

$$\langle x_1 | e^{-iH(t_1 - t_0)} | x_0 \rangle. \tag{1.28}$$

We could insert spurious $t_{1/N}$, $t_{2/N}$,... $t_{(N-1)/N}$ and decompose

$$e^{-iH(t_1-t_0)} = \prod_{j=1}^{N} e^{-iH(t_{j/N} - t_{(j-1)/N})} \tag{1.29}$$

and we can also insert a superfluous complete set of states $|x_{j/N}\rangle$ corresponding to intermediate positions, at time $t_{j/N}$, satisfying the completeness relation

$$\int_{\infty}^{\infty} dx_{j/N} |x_{j/N}\rangle\langle x_{j/N}| = \mathbb{I}. \tag{1.30}$$

In this way, the amplitude is

$$\int_{\mathbb{R}^{N-1}} \prod_{j=1}^{N-1} dx_{j/N} \ \prod_{j=1}^{N} \langle x_{j/N} | e^{-iH(t_{j/N} - t_{(j-1)/N})} | x_{(j-1)/N} \rangle. \tag{1.31}$$

To simplify the situation, say that $t_{j/N} - t_{(j-1)/N}$ is constant, Δt, i.e. $t_{j/N} = t_0 + j\Delta t$ with $\Delta t = (t_1 - t_0)/N$. If N is large, then Δt is small and the first order approximation in Baker-Campbell-Hausdorff formula, gives

$$\langle x_{j/N} | e^{-iH\Delta t} | x_{(j-1)/N} \rangle \sim \langle x_{j/N} | e^{-i\Delta t p^2/2m} | x_{(j-1)/N} \rangle e^{-i\Delta t V(x_{(j-1)/N})}. \tag{1.32}$$

On the other hand, in the moment space, $\langle x_{j/N} | e^{-i\Delta t p^2/2m} | x_{(j-1)/N} \rangle$ equals

$$\int \frac{dp}{2\pi} e^{-i\Delta t p^2/2m} \langle x_{j/N} | p \rangle \langle p | x_{(j-1)/N} \rangle = \int \frac{dp}{2\pi} e^{-i\Delta t p^2/2m + ip(x_{j/N} - x_{(j-1)/N})}. \tag{1.33}$$

Computing the integral [39] and substituting in (1.31), we have that

$$\langle x_1 | e^{-iH(t_1-t_0)} | x_0 \rangle \sim \int_{\mathbb{R}^{N-1}} \mathcal{D}_N x \ e^{i\mathcal{S}_N} \tag{1.34}$$

where we are using the abbreviations

$$\mathcal{D}_N x = \frac{\prod_{j=1}^{N-1} dx_{j/N}}{(2\pi i \Delta t/m)^{n/2}} \quad \text{and} \quad \mathcal{S}_N = \Delta t \sum_{j=1}^{N} \Big(\frac{m}{2} \Big(\frac{x_{j/N} - x_{(j-1)/N}}{\Delta t} \Big)^2 - V(x_{(j-1)/N}) \Big). \tag{1.35}$$

One can consider $\mathcal{S}_N$ as a Riemann sum of the classic action $S = \int_0^{t_1 - t_0} L \ dt$ and we could interpret that the measure $\mathcal{D}_N x$ in the limit is a measure $\mathcal{D}x$ in the infinite dimensional space of all possible paths. The approximation (1.32) via Baker-Campbell-Hausdorff formula becomes equality when $N \to \infty$, then we have the appealing formula

$$\langle x_1 | e^{-iHT} | x_0 \rangle = \int \mathcal{D}x \ e^{i\int_0^T L \ dt}. \tag{1.36}$$

In our case, $L = \frac{1}{2}m\dot{x}^2 - V$. This is a *path integral*. The suggestive and interesting point is that if we blindly apply the *stationary phase principle* we have the classical trajectories via the *principle of least action*.

The quantum field analog of (1.36) is that the vacuum expectation of an observable O is given by

$$\langle O \rangle = \frac{1}{Z}\int \mathcal{D}\phi\, O e^{iS[\phi]} \qquad \text{with} \quad Z = \int \mathcal{D}\phi\, e^{iS[\phi]}. \tag{1.37}$$

Although (1.36) is very impressive as well the derivation of classical trajectories, in a first reading, the path integral formalism is nothing else that an alternative approach to write the solutions of Schrödinger's equation. In fact, in [23] is presented in this way (see the "Purpose of the book" in p.23). There is also a pitfall. It is not clear if the definition is sound because the convergence of the sequence of the finite dimensional highly oscillatory integrals is doubtful.

Part of the calculations of path integrals depend on the following formula, note the analogy with (1.37), valid for any vector $\vec{b} \in \mathbb{R}^N$ and a positive $N \times N$ matrix A

$$\frac{I_N(A, \vec{b})}{I_N(A, \vec{0})} = e^{\frac{1}{2}\vec{b}^t A^{-1}\vec{b}} \qquad \text{when} \quad I_N(A, \vec{b}) = \int e^{-\frac{1}{2}\vec{x}^t A\vec{x} + \vec{b}\cdot\vec{x}} d^N x. \tag{1.38}$$

The form of the result does not depend on the dimension. When studying scalar free fields in quantum field theory, in the Lagrangian we find something like $\partial^\mu\partial_\mu + m^2$ instead of A and a field φ instead of $\vec{x}$ and one hopes that the formula is valid replacing A^{-1} by the corresponding inverse operator, closely related to the *Green function*, the *propagator*. This is very appealing but, even overlooking the positivity, clearly the Lagrangians have not in general a quadratic-like form[4]. One may conjecture a perturbative quadratic approximation can be done paying with some polynomial terms like in the real numbers $\int e^{-x^2+\lambda x^n} dx \sim \int (1 + \lambda x^n) e^{-x^2} dx$ for small λ. The different contributions of these polynomial terms can be counted by pictorial representations, the *Feynman diagrams*[5].

Even without a mathematical sound definition, the path integral gives a deep insight about the simplest cases and about the perturbative regime in general.

[4]In [79, I.11], A. Zee makes ironic but illustrative comments: "Quantum field theory is not that difficult; it just consists of doing one great big integral [...] Quantum field theorists try to dream up ways to evaluate (1) [the path integral for the ϕ^4 Lagrangian], and failing that, they invent tricks and methods for extracting the physics they are interested in, by hook and by crook, without actually evaluating (1)."

[5]In part they were anticipated by E. Stueckelberg, a physicist with several unrecognized contributions to quantum field theory.

The lattice approach to the non perturbative theory

Browsing books of quantum field theory, one might conclude that the fundamental part of the subject consists of drawing, understanding and computing Feynman diagrams. The sad truth is that this emphasis depends on our poor ability to do calculations in the non perturbative regime. On one hand Feynman trumpets the precision of the theoretical calculations of the magnetic moment of the electron in 1983 ("If you were to measure the distance from Los Angeles to New York to this accuracy, it would be exact to the thickness of a human hair" [22]) and on the other hand, even today, we have only mild theoretical estimations for the same quantity for the proton. In highly non perturbative theories like the strong interaction and in general in Yang-Mills theories, we are losing the big picture (but something can be saved [55]). Even leaving aside the calculations, from the theoretical point of view it seems that destroying the high symmetry of Yang-Mills theory, the gauge invariance, is a wrong approach.

In this context, and with the increasing capabilities of the computers, *lattice gauge theory* has acquired a growing interest. It was introduced by K.G. Wilson in [75]. The idea is using a discretization of the spacetime, through a lattice $(a\mathbb{Z})^4$ and, quoting Wilson, "a lattice version of Euclidean vacuum expectation values, starting from a lattice version of the Feynman path integral". The value of a^{-1} plays the role of an ultraviolet cutoff that breaks Lorentz covariance but, on the other hand, there is a perfect preservation of the symmetries of the Lagrangian. Zee [79, VII.1] jokes "Wilson proposed a way out: Do violence to Lorentz invariance rather to gauge invariance".

In the next lines we briefly review the elements of lattice gauge theory and we show how to recover (invent?) the Yang-Mills Lagrangian from this approach. For more information see the monographs [61] [57].

The analogue of (1.37) on the lattice is

$$\langle O\rangle = \frac{1}{\mathcal{Z}}\int \prod_{x,\mu} dU_\mu(x)\; O e^{-S_W} \qquad \text{with} \quad \mathcal{Z} = \int \prod_{x,\mu} dU_\mu(x)\; e^{-S_W}. \tag{1.39}$$

It is apparent that it has been introduced a Wick rotation to regularize (see [57, p.21]) that makes the approach closer to statistical mechanics [41]. The spacetime becomes Euclidean.

Let us explain the ingredients in (1.39).

The points x are points in the lattice, i.e. of the form $x = an$ with $n \in \mathbb{Z}^4$. A straight line path joining two neighbors of the lattice in the $\hat\mu$ direction, x and $x + \hat\mu$ is called a *link*. A *plaquette* $P(x;\mu,\nu)$ is the square bounded by 4 links with vertexes x, $x + a\hat\mu$, $x + a\hat\mu + a\hat\nu$ and $x + a\hat\nu$ with $\mu \neq \nu$.

For each link connecting x and $x + a\hat{\mu}$ we have a *parallel transporter* $U_\mu(x)$ that takes the field from one extreme of the link to another. Under gauge transformations, it follows the rule.

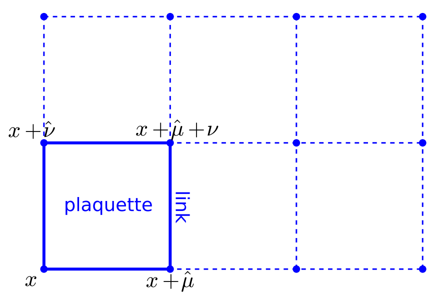
$$U_\mu(x) \longmapsto \Omega(x)U_\mu(x)\Omega(x + a\hat{\mu}). \tag{1.40}$$

For a closed path C on the lattice, a *Wilson loop*, the trace of the product of all parallel transporters in the path is gauge invariant.

When we do this in a plaquette $P(x;\mu,\nu)$ we have a natural gauge invariant action

(1.41)
$$S_{P(x;\mu,\nu)} = \mathrm{Tr}\big(\mathcal{U}(x;\mu,\nu)\big) \qquad \text{with} \quad \mathcal{U}(x;\mu,\nu) = U_\mu(x)U_\nu(x+a\hat{\mu})U_\mu^\dagger(x+a\hat{\nu})U_\nu^\dagger(x).$$

The action S_W appearing in (1.39) essentially is the sum of $S_{P(x;\mu,\nu)}$ over all plaquettes with a normalization constant (see [48] [41]).



In the plaquette

$$A = U_\mu(x)U_\nu(x+a\hat{\mu})U_\mu^\dagger(x+a\hat{\nu})U_\nu^\dagger(x)$$
$$\downarrow$$
$$\Omega(x)A\Omega^\dagger(x)$$

Finally, the (finite) integration in (1.39) is over all possible parallel transporters. For the gauge group that we consider, over all possible matrices in $SU(N)$. In this compact group there is an invariant measure, the *Haar measure*, that is the one employed to integrate [25].

Noticeable limiting cases

We are going to check that the Yang-Mills action is recovered if we assume that the parallel transport is done through $\mathrm{Pexp}\big(-i\int_\gamma A_\mu dx^\mu\big)$ where Pexp denotes the *path ordered integral* (by simplicity we omit the coupling constant). The motivation of this operator[6] in quantum mechanics and in QED can be found in the old paper [21].

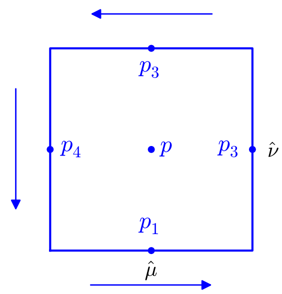
For a small this gives $U_\mu(x) \approx e^{-iaA_\mu(x+a\hat{\mu}/2)}$ and substituting in (1.41)

$$S_{P(x;\mu,\nu)} = \mathrm{Tr}\Big(e^{-iaA_\mu(p_1)}e^{-iaA_\nu(p_2)}e^{iaA_\mu(p_3)}e^{iaA_\nu(p_4)}\Big) \tag{1.42}$$

where the notation is:

[6] Mathematically is as simple as saying that to solve the matrix differential equation $X'(t) = AX(t)$, $X(0) = \mathbb{I}$ one needs the exponential, $X(t) = e^{tA}$ but if A depends on t, one needs the path (time) ordered integral (see also [25, p.555]).

$$p = x + \frac{a}{2}(\hat{\mu} + \hat{\nu}) \qquad \begin{aligned} p_1 &= x + \frac{a}{2}\hat{\mu} \\ p_2 &= x + a\hat{\mu} + \frac{a}{2}\hat{\nu} \\ p_3 &= x + \frac{a}{2}\hat{\mu} + a\hat{\nu} \\ p_4 &= x + \frac{a}{2}\hat{\nu} \end{aligned}$$



By the Baker-Campbell-Hausdorff formula, to order 2 the product of the two first exponentials amounts

$$\exp\Big(- iaA_\mu(p_1) - iaA_\nu(p_2) - \frac{1}{2}a^2[A_\mu(p), A_\nu(p)]\Big). \tag{1.43}$$

Taylor expanding around p,

$$\exp\Big(- ia\big(A_\mu - \frac{a}{2}\partial_\nu A_\mu\big) - ia\big(A_\nu + \frac{a}{2}\partial_\mu A_\nu\big) - \frac{1}{2}a^2[A_\mu, A_\nu]\Big) \tag{1.44}$$

where the field and its derivatives are evaluated at p. In the same way, the product of the last exponentials gives

$$\exp\Big(ia\big(A_\mu + \frac{a}{2}\partial_\nu A_\mu\big)ia\big(A_\nu - \frac{a}{2}\partial_\mu A_\nu\big) - \frac{1}{2}a^2[A_\mu, A_\nu]\Big). \tag{1.45}$$

Now, we multiply (1.44) and (1.45) to get via (1.42)

$$S_{P(x;\mu,\nu)} \sim \mathrm{Tr}\Big(\exp\big(ia^2\partial_\nu A_\mu - ia^2\partial_\mu A_\nu + a^2[A_\mu, A_\nu]\big)\Big). \tag{1.46}$$

Recalling the definition of the curvature (1.24), we have

$$S_{P(x;\mu,\nu)} \sim \mathrm{Tr}\Big(\exp\big(- ia^2F_{\mu\nu}\big)\Big) \sim 1 - \frac{1}{2}a^4\mathrm{Tr}\big(F_{\mu\nu}\big)^2. \tag{1.47}$$

Then $a^{-4}\big(S_{P(x;\mu,\nu)} - 1\big)$ is like the (Euclidean) Yang-Mills Lagrangian (1.25) and when we sum over the plaquettes we obtain the lattice approximation of the Yang-Mills action.

Wilson entitled his paper [75] "Confinement of quarks". A fascinating aspect of lattice gauge theory is that gives some hint about the, still unproved, confinement properties of Yang-Mills theories. Without entering into details (see [48] for short simple explanations), the point is that in the strong coupling limit, when the coupling constant goes to ∞, we can do some natural approximations. Consider a rectangular Wilson loop C, the boundary of a rectangle of dimensions R and T, where R physically represents the separation between quarks and T is the time. If $W(C)$ is the (normalized) trace of the product of the parallel transporters corresponding to C, then it can be proved that the vacuum expectation verifies, with suitable normalizations, $\log\langle W(C)\rangle \sim -VR$ for T large where V is the potential energy. On the other hand, the strong coupling limit suggests an *area law* $\log\langle W(C)\rangle \sim -K\text{Area}(C)$. Thus one would obtain that the potential energy is proportional to the distance, than would imply confinement.

1.3 Large N

Motivation and results

In 1974, shortly before [75], 't Hooft introduced the use of the number of colors N as a large parameter in non Abelian gauge theories [64]. In order to get a nontrivial limit theory, one has to run accordingly the coupling constant in such a way that

$$g^2N \to \text{ constant } \asymp 1 \quad \text{whenever } N \to \infty. \tag{1.48}$$

This is the *'t Hooft coupling* (not the only possibility [14]). See the reasons for the numerology in [79] and [48].

Perhaps the closer analogy is in statistical mechanics where expanding in a large unknown number of particles turns out natural. Some authors have also pointed the analogy with [63] where an approximation to a model of phase transition is solved assuming that the spin goes to infinity.

In principle it seems an idea very far from physical reality because in the Standard Model we have a strong force with just three colors but, quoting 't Hooft, "The $1/N$ expansion may be a reasonable perturbation expansion, in spite of the fact that N is not very big." [64]. Here we are going to emphasize two point in favor of these expansions. The first one is phenomenology and the second and most important point is the simplification of the theories.

After [64], 't Hooft published a model for mesons based on it [65]. Years later, E. Witten wrote about baryons that are more complicated [76]. Although these theories do not seem very physical (it is a $1+1$ dimensional theory for mesons and baryons have N quarks), there are some noticeable results assuming confinement at $N=\infty$ suggesting that we are under the right phenomenology.

We simply quote here some of the items listed in [14]. See [11] and [76] for more information.

- The decay amplitude of mesons is $O\big(N^{-1/2}\big)$ (stability).
- The scattering amplitude of mesons is $O\big(N^{-1}\big)$ (no interaction).
- The vertexes corresponding to l glueballs contribute $O\big(N^{1-l}\big)$ (no interaction glueball-glueball).

- The vertexes corresponding to l glueballs and k mesons contribute $O\big(N^{1-l-k/2}\big)$ (no interaction glueball-meson).
- Baryon masses are $O(N)$.
- The typical baryon-baryon vertex contributes $O(N)$ (strong interaction between baryons).

According to [11] "For mesons, things are wonderful. [...] the properties we found [...] form an accurate portrait of the mesons. They form a caricature. But it is a recognizable caricature; [...] For the baryons, things are not so good."

Beyond these and other phenomenological "caricatures", probably the most important reason to consider large N is that it simplifies the theories. Yang-Mills models are so complicated that researchers are eager for finding toy models to play with. One of the simplifying features is that in many interactions the large N limit induces *factorization* of gauge invariant operators. This is a key point for the *volume independence* (reduction) in some models that will appear in the next chapter. The factorization is essentially the independence that allows to write the vacuum expectation value of a product of traced observables as the product of vacuum expectations.

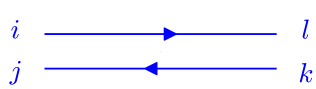
Planar diagrams

Let us see the kind of simplification of the theory that 't Hooft got in the original paper [64] in the large N limit. He considered $U(N)$ as the gauge group instead of $SU(N)$. This is simply a technical point. With the usual choice of the normalization, the generators T^a of the Lie algebra of $SU(N)$ satisfy

$$\sum_a T^a_{ij} T^a_{kl} = \frac{1}{2}\delta_{il}\delta_{jk} - \frac{1}{2N}\delta_{ij}\delta_{kl}. \tag{1.49}$$

In $U(N)$ the last term does not appear but for large N if we only consider leading terms, both cases are equivalent.

With the *double line notation* [48], each gluon propagator is represented by a double line contributing as in (1.49)



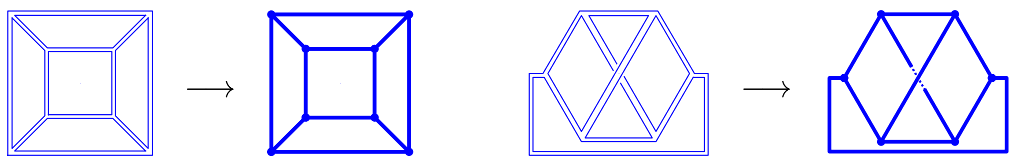
$$\propto \quad \delta_{il}\delta_{jk}$$

Since there are N colors, each loop gives a contribution proportional to N. On the other hand, the 3-gluon vertex and the 4-gluon vertex give contributions proportional to g and g^2 in their diagrams. Then the diagrams are associated with a factor

$$g^{V_3+2V_4} N^I \qquad \text{where} \quad \begin{cases} I = \# \text{ “index” (gluon) loops,} \\ V_3 = \# \text{ 3-gluon vertexes,} \\ V_4 = \# \text{ 4-gluon vertexes.} \end{cases} \tag{1.50}$$

To ease the exposition, we assume, as in [70], that there are not external lines (we deal with *vacuum diagrams*) and there are not quark loops. The first assumption can be dropped introducing 2-vertexes and the second with a global factor (see the details in [70]).

Drawing thick lines instead of double lines we can transform the diagrams into graphs (by simplicity we do not draw arrows)

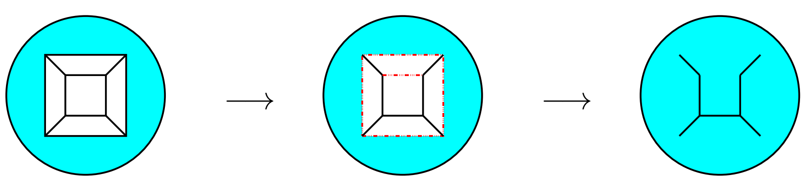


With the usual notations, let V, E and F be the number of vertexes, edges and faces[7] of the graph (including the exterior one). We know that $2E = \sum_j jV_j$ because each edge connects two vertexes. Then (1.50) is

$$g^{v_3+2V_4} N^I = g^{2E-2V} N^F = (g^2 N)^{F-\chi} N^\chi \tag{1.51}$$

where $V - E + F = \chi$ is the *Euler characteristic*, a topological invariant that reaches its maximal value $\chi = 2$ for planar graphs, those that can be drawn on the plane without self-intersections. Then in the large N limit (1.48), *planar diagrams* are dominant with a contribution $\propto N^2$. This is a huge reduction on the diagrams to be considered, and shows the kind of vast simplifications that appear in the large N limit. In the figures above, the first diagram is planar and contributes $g^8 N^6 = \lambda^4 N^2$ with $\lambda = g^2 N$ and the second diagram is not planar and contributes $g^6 N^3 = \lambda^3$.

To deduce that for every planar graph $V - E + F = 2$, there is a simple intuitive argument [56, §12]. Just consider the planar graph as an island with $F - 1$ fields. If we break a dyke to flood each field (edge), we reduce the problem to the study of the trivial case of a tree ($f - 1 = 0$).



Non planar graphs are drawn on surfaces with handles that connect fields and allow to perform the flooding more efficiently. It explains why χ is smaller.

[7]Strictly speaking, the definition of each face is associated to the embedding of the graph on a surface, rather than to the graph itself.

A comment about large N simulations

Doing numerical simulations of large N theories is not an easy task. Assume for simplicity that in (1.39) we are only considering a finite number of values of x (see the models of the next chapter). Although $\mathcal{Z}$ is a finite integral, is too complicate to try deterministic integrators. The natural approach is Monte Carlo integration. A key point is to generate random $SU(N)$ matrices with a certain distribution [13].

Let us think firstly in a simpler problem: Simulate a sample of a real random variable (in $\mathbb{R}$) with density function compactly supported in $[a, b]$. A useful general purpose method is *rejection sampling*. It reduces to the single steps:

- Generate (x, y) under an uniform distribution in $[a, b] \times [0, \max f]$.
- If $y < f(x)$ accept x in the sample.
- Repeat the process until having a sample of the required cardinality.

This works quite well (although there are better algorithms in special cases) but if we try to adapt it to $SU(N)$ with N large, we shall suffer what is called in several areas of statistics the *curse of dimensionality*.

We just mention here the ingenious idea introduced in [8] and [52] that is the basis of several approaches. To generate matrices of $SU(N)$, one considers

$$(1.52)\qquad \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & A_{N-1} \end{pmatrix} \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & A_{N-2} & \\ & & & & 1 \end{pmatrix} \cdots \begin{pmatrix} A_1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix}$$

where A_1, A_2,..., A_{N-1} are matrices of $SU(2)$, that have low dimension. Thanks to the properties of the Haar measure, the A_j's inherits in some way the required distribution for $SU(N)$.

See also [15] for an overview of the current methods and [28] focused on the kind of models that we shall treat later.

Chapter 2

A twisted model on the flat 2-torus

2.1 Brief review on the EK and the TEK models

Volume independence

In some models the combination of lattice gauge theory and large N induces an interesting property called *volume independence* (or *reduction*). These models are important by themselves but we shall be extremely sketchy here because we do not need them for our purposes. Consequently in this section we just mention some aspects of the models as a motivation without entering into details. For more information see [7, Ch.7], [47, §4] and [41].

If we recall the partition function $\mathcal{Z}$ in lattice gauge theory (1.39), we note that the product extends to the, in principle, infinitely many points in the lattice. This is unaffordable from a practical point of view and it suggest to impose periodic conditions. We shall come back on this in the next section but we should keep in mind this framework now.

T. Eguchi and H. Kawai proved in [18] a surprising result assuming certain facts: In the large N limit, $\mathcal{Z}$ in (1.39) can be replaced by

$$\mathcal{Z}_{EK} = \int \prod_{\mu} dU_\mu(x)\, e^{-S_{EK}}. \tag{2.1}$$

where, up to normalizing factors, S_{EK} is

$$\sum_{\mu\neq\nu} \mathrm{Tr}\big(U_\mu U_\nu U_\mu^\dagger U_\nu^\dagger\big). \tag{2.2}$$

If we compare it with (1.39) and (1.41), we see that the dependance on the points of the lattice has disappeared. This is the aforementioned volume independence (or reduction, for the reduction on the degrees of freedom) and the resulting model is the *Eguchi-Kawai model* (EK). The authors mention in their paper that the result is based on three assumptions. Two of them have to do with the loop equation and its relation with factorization, and the third one is related to symmetry. We briefly review them here.

Loop equations and factorization

To establish the equivalence between the model on the lattice and the reduced one, identifying all the $U_\mu(x)$ for a given μ into a single U_μ, Eguchi and Kawai considered Wilson loops and proved that they satisfy the same kind of equations in both models. These are the *loop equations*. For a simple loop C (without self-intersections) they read

$$\big\langle \mathrm{Tr}\big(W(C)\big)\big\rangle \propto \sum_{\mu\neq\nu}\Big(\big\langle \mathrm{Tr}\big(W(C)\mathcal{U}(x;\mu,\nu)\big)\big\rangle - \big\langle \mathrm{Tr}\big(W(C)\mathcal{U}^{\dagger}(x;\mu,\nu)\big)\big\rangle\Big) \tag{2.3}$$

with a normalizing constant coming from the action, where $\mathcal{U}(x;\mu,\nu)$ is like in (1.41) and $W(C)$ is the product of the parallel transporters along C. If C is not simple, say that it can be divided into two simple paths C_1 and C_2, then an extra term $\big\langle \mathrm{Tr}\big(W(C_1)\big)\mathrm{Tr}\big(W(C_2)\big)\big\rangle$ appears. Assuming factorization at $N=\infty$, it gives $\big\langle \mathrm{Tr}\big(W(C_1)\big)\big\rangle\big\langle \mathrm{Tr}\big(W(C_2)\big)\big\rangle$.

Symmetry breaking and the twisted model

After constructing the loop equations, the argument in [18] depends on canceling some traces corresponding to open paths. Algebraically, it reflects an inner symmetry: Everything should be invariant under the change $U_\mu \mapsto e^{2\pi ik/N}U_\mu$ because this transformation preserves $SU(N)$ and the constant factors commute with any matrix (they are in the center of the group via $\lambda \mapsto \lambda\mathbb{I}$). If a path is open then for certain μ there are U_μ not compensated by $U_\mu^{\dagger}$ and the symmetry proves that the trace vanishes.

After the publication of [18] (in fact the reference appears as a note added in proof), it was shown [4] that this symmetry is spontaneously broken for the weak coupling. Nevertheless the property of volume independence is so appealing that it is worthy to look for modifications in the original model.

The main modification was introduced shortly after by A. González-Arroyo and M. Okawa in [34] and [33]. It is called the *twisted Eguchi-Kawai model* (TEK), The novelty is a change in the action introducing an N-root of the unity that comes from the generalization of the periodic boundary conditions and is motivated by the twists introduced in [66] in other context. The possible symmetry breaking in the modified model is discussed in [71], [35] and [37]. An interesting point is that the choice of the twist gives more chances to preserve the symmetry.

2.2 Yang-Mills theory on a twisted torus

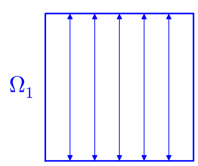
Definition of the bundle

The configuration space of the model is $\mathbb{T}^2$, the flat torus with periods L and L. Formally it is the quotient space

$$\mathbb{T}^2 = \mathbb{R}^2/(L\mathbb{Z})^2 \qquad \text{endowed with the metric} \quad dx^2 + dy^2, \tag{2.4}$$

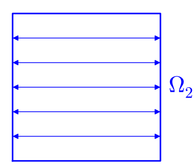
where $(L\mathbb{Z})^2$ identifies the pairs (x_1, x_2) and $(x_1 + n_1 L, x_2 + n_2 L)$ with $n_1, n_2 \in \mathbb{Z}$.

We consider an $SU(N)$ principal fiber bundle to define the gauge group. Then the vector potentials are represented by A_j, $j = 1, 2$ (corresponding to each coordinate) and belonging to $\mathfrak{su}(N)$, i.e. they are $N \times N$ traceless hermitian matrices. In principle one is tempted to consider them as functions of each point of $\mathbb{T}^2$ or, equivalently, defined in $\mathbb{R}^2$ imposing that A_j is L periodic in each variable. But this is too demanding from the physical point of view because the potentials are only defined modulo gauge transformations. From the mathematical point of view, one must define *transition functions* (matrices in this case) to determine the bundle and the connection (the potential) must be invariant by them. Hence, we consider $\Omega_j(x)$, $j = 1, 2$ specifying how to glue the horizontal and the vertical boundaries in the square representation.



$$\Omega_1(x) = \Omega_1(x_2)$$

$$\Omega_2(x) = \Omega_2(x_1)$$



Then the natural condition for the gauge fields to be consistent with the topology of the bundle is that $A_j(x + L\vec{e}_l)$ is the same as $A_j(x)$ after applying the gauge transformation $\Omega_l(x)$, where $\{\vec{e}_1, \vec{e}_2\}$ is the usual canonical basis of $\mathbb{R}^2$. Recalling (1.22), we have

$$A_j(x + L\vec{e}_l) = \Omega_l(x) A_j(x) \Omega_l^\dagger(x) + i\Omega_l(x)\partial_j \Omega_l^\dagger(x). \tag{2.5}$$

Note that $\Omega^\dagger = \Omega^{-1}$ in $SU(N)$. We have a lot of freedom to choose the Ω_l's but they are not independent functions when considered on $\mathbb{R}^2$ because (2.4) requires the commutation relation $L\vec{e}_j + L\vec{e}_l = L\vec{e}_l + L\vec{e}_j$, that characterizes the Abelian group $(L\mathbb{Z})^2$,to be preserved. Then the actions of the gauge transformations $\Omega_j(x + L\vec{e}_l)\Omega_l(x)$ and $\Omega_l(x + L\vec{e}_j)\Omega_j(x)$ must coincide acting on vector potentials. In principle this gives complicate equations using (2.5), but they simply reduce to say that

$$\Omega_j(x + L\vec{e}_l)\Omega_l(x)\big(\Omega_l(x + L\vec{e}_j)\Omega_j(x)\big)^\dagger \tag{2.6}$$

commutes with any gauge potential. One could take it as the identity $\mathbb{I}$, this is the *no-twist condition*, but it could be in general of the form $\lambda\mathbb{I}$ with $\lambda \in \mathbb{C}$. Noting that the determinant implies $\lambda^N = 1$ and that trivially $\lambda = 1$ for $j = l$, the most general consistency conditions are the *twisted boundary conditions* (cf. [67])

$$\Omega_j(x + L\vec{e}_l)\Omega_l(x) = e^{2\pi i n_{jl}/N}\Omega_l(x + L\vec{e}_j)\Omega_j(x) \qquad \text{where} \quad n_{jl} = -n_{lj}. \tag{2.7}$$

This establishes different topological *twist sectors* according to the choice of n_{12}.

Gauge transformations are actually defined modulo the center of the gauge group (meaning that the center acts trivially). This is the *center symmetry*[1]. The value of $\Omega_j(x)\Omega(x)\Omega_j(x)^\dagger$could differ from $\Omega(x + L\vec{e}_j)$ in an element of the center, and we then have

$$\Omega(x + L\vec{e}_j) = e^{2\pi i k_j/N}\Omega_j(x)\Omega(x)\Omega_j(x)^\dagger. \tag{2.8}$$

As mentioned in [32], n_{jk} measures the topological obstruction to lift the bundle from $SU(N)/Z_N$ to $SU(N)$.

To avoid problems with subgroups, from now on we assume that N is a prime number $N > 2$. The rough idea is that if N is composite then the center is a direct product of smaller groups that would give a more involved sector structure.

We have a lot of freedom to choose valid transition matrices, i.e. satisfying (2.7). Since we are going to use them in some way to analyze the gauge field, it is convenient to do it in a simple but still versatile way. A natural one is taking constant matrices. This kind of solutions are called *twist eaters* and they also appear in the lattice gauge approach [41].

In the papers [73] and [45], it has been studied the problem of finding the most general formulas for twist eaters not related by similarity transformations (global gauge transformations). Here we give a simple proof that fits our case. If we call Γ_1 and Γ_2 the constant matrix solutions of (2.7), then

$$\Gamma_2 = e^{2\pi i n_{12}/N}\Gamma_1^{-1}\Gamma_2\Gamma_1. \tag{2.9}$$

Hence the eigenvalue set is invariant under multiplication by $e^{2\pi i n_{12}/N}$. We leave apart the no-twist condition ($n_{12} = 0$). Then, the eigenvalue set is also invariant by $e^{2\pi i/N}$. Note that this requires the absence of proper subgroups of $\mathbb{Z}_N$ or equivalently that N is prime. Then with a similarity transformation, the one changing to the Jordan basis, we have that

$$\Gamma_2 = \xi^{(1-N)/2}\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & \xi & 0 & \dots & 0 \\ 0 & 0 & \xi^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \xi^{N-1} \end{pmatrix} \qquad \text{with} \qquad \xi = e^{2\pi i/N}. \tag{2.10}$$

[1]In [66] it is considered that the actual gauge group is $SU(N)/Z$.

The outer factor $\xi^{(1-N)/2}$ is just to force $\Gamma_2 \in SU(N)$. Substituting in (2.9) and multiplying by $\xi^{(N-1)/2}\Gamma_1$ in both sides, we have that the element jk of Γ_1, say $\big(\Gamma_1\big)_{jk}$, satisfies

$$\big(\Gamma_1\big)_{jk} = e^{2\pi i(n_{12}+k-j)/N}\big(\Gamma_1\big)_{jk}. \tag{2.11}$$

Then we have the natural choice of the permutation matrix corresponding to an n_{12}-shift modulo N. If $n_{12} > 0$ (which we can assume interchanging Γ_1 and Γ_2)

$$\big(\Gamma_1\big)_{jk} = \begin{cases} 1 & \text{if } 1 \le j = k + n_{12} - N \le n_{12}, \\ 1 & \text{if } 1 \le k = j - n_{12} \le N - n_{12}, \\ 0 & \text{otherwise.} \end{cases} \tag{2.12}$$

With any constant matrix choice, (2.5) reads

$$A_j(x + L\vec{e}_l) = \Gamma_l A_j(x)\Gamma_l^\dagger. \tag{2.13}$$

For each $\vec{e} = (e_1, e_2)$ defined modulo N, we consider the $N \times N$ matrices (cf. [68])

$$\widehat{\Gamma}(\vec{e}) = \frac{1}{\sqrt{2N}} e^{i\alpha}\Gamma_1^{e_1}\Gamma_2^{e_2} \tag{2.14}$$

where α is an arbitrary phase to be fixed later. These vectors can be considered as representations of the center symmetry, associated to the group $\mathbb{Z}_N \times \mathbb{Z}_N$, one cyclic group per coordinate. They are called (chromo-)*electric flux vectors*.

It holds

$$\Gamma_j\widehat{\Gamma}(\vec{e})\Gamma_j^\dagger = e^{2\pi i n_{jk}e_k/N}\widehat{\Gamma}(\vec{e}) \tag{2.15}$$

where in $n_{jk}e_k$ is assumed summation on k. Recall that n_{jk} was skew-symmetric (2.7), then there is only a term in the sum. The relation (2.15) follows from (2.7) using the identities

$$\Gamma_1\big(\Gamma_1^{e_1}\Gamma_2^{e_2}\big)\Gamma_1^{-1} = \Gamma_1^{e_1}\big(\Gamma_1\Gamma_2\Gamma_1^{-1}\big)^{e_2} \qquad \text{and} \qquad \Gamma_2\big(\Gamma_1^{e_1}\Gamma_2^{e_2}\big)\Gamma_2^{-1} = \big(\Gamma_2\Gamma_1\Gamma_2^{-1}\big)^{e_1}\Gamma_2^{e_2}. \tag{2.16}$$

Expansions of the gauge field

Taking traces in (2.15), we have that for $\vec{e} \neq \vec{0}$ modulo N then the matrices $\widehat{\Gamma}(\vec{e})$ are traceless. These $N^2 - 1$ matrices are linearly independent as the explicit formulas (2.12) and (2.10) show. We can express the gauge fields as a linear combination of these matrices

$$A_j(x) = \sum_{\vec{e}} \lambda_j(x, \vec{e})\widehat{\Gamma}(\vec{e}) \qquad \text{with} \qquad \lambda_j(x, \vec{e}) \in \mathbb{C}. \tag{2.17}$$

Here $\vec{e}$ runs over a complete reduced class of residue classes of vectors modulo N, i.e. all the residue classes except that of $\vec{0}$. We assume this restrictions in all the sums involving $\vec{e}$. Using (2.13), (2.15) and the uniqueness of the λ_j's (by linear independence),

$$\lambda_j(x + L\vec{e}_l, \vec{e}) = \lambda_j(x, \vec{e}) e^{2\pi i n_{lk} e_k / N}. \tag{2.18}$$

In particular, $\lambda_j(x, \vec{e}) e^{-2\pi i x_l n_{lk} e_k / LN}$, with summation on $l, k \in \{1, 2\}$, is a L periodic function in x_1 and x_2. We can hence expand it into Fourier series with orthonormal harmonics $L^{-1}\exp(2\pi i m_j x_j / L)$ [72]. Substituting this expansion, we can re-write (2.17) in the form

$$A_j(x) = \frac{1}{L} \sum_{\vec{p}} \hat{A}_j(\vec{p}) e^{i\vec{p}\cdot\vec{x}} \widehat{\Gamma}(\vec{e}) \qquad \text{with} \quad \hat{A}_j(\vec{p}) \in \mathbb{C}, \tag{2.19}$$

where the notation is as follows:

$$\vec{x} = (x_1, x_2), \quad \vec{p} = \vec{p}^{(s)} + \vec{p}^{(c)}, \quad p_j^{(s)} = \frac{2\pi m_j}{L}, \quad p_j^{(c)} = \frac{2\pi n_{jk} e_k}{LN}. \tag{2.20}$$

Following [29], we can say that $\vec{p}^{(c)}$ is the *color-momentum*, coming from the decomposition (2.17) of the gluon field, and $\vec{p}^{(s)}$ is the *spatial-momentum*, the usual one in quantum mechanics: The variable in the (discrete) Fourier transform corresponding to the space.

The total momentum is then of the form

$$\vec{p} = \frac{2\pi\vec{n}}{LN} \quad \text{with } \vec{n} \in \mathbb{Z}^2 \qquad \text{and} \qquad \vec{n} \equiv n_{12}(e_2, -e_1) \pmod{N}. \tag{2.21}$$

Many values of $\vec{p}$ correspond to the same electric flux vector $\vec{e}$ and hence to the same $\widehat{\Gamma}(\vec{e})$ in (2.19). We are going to let the arbitrary phase in (2.14) to depend on $\vec{p}$. With a suitable choice, namely (see [29])

$$\alpha(\vec{p}) = \frac{\overline{n}_{12} N L^2}{4\pi} p_1 p_2 \qquad \text{with} \quad \overline{n}_{12} n_{12} \equiv 1 \pmod{N}, \tag{2.22}$$

we impose

$$\widehat{\Gamma}(-\vec{p}) = \widehat{\Gamma}^\dagger(\vec{p}) \tag{2.23}$$

where $\widehat{\Gamma}(\vec{p})$ stands for $\widehat{\Gamma}(\vec{e})$ with the indicated choice of α. Then (2.19) reads

$$A_j(x) = \frac{1}{L} \sum_{\vec{p}} \hat{A}_j(\vec{p}) e^{i\vec{p}\cdot\vec{x}} \widehat{\Gamma}(\vec{p}). \tag{2.24}$$

Multiplying by $\widehat{\Gamma}(-\vec{q})e^{-i\vec{q}\cdot\vec{x}}$, taking into account (2.14) and (2.23), we have the Fourier coefficient formula (cf. [72])

$$\hat{A}_j(\vec{p}) = \frac{2}{L}\iint_{\mathbb{T}} \mathrm{Tr}\big(\widehat{\Gamma}(-\vec{p})A_j(x)\big)e^{-i\vec{p}\cdot\vec{x}}d^2\vec{x}. \tag{2.25}$$

Note that the 2 factor comes from $\mathrm{Tr}(\mathbb{I}) = N$ times the normalizing factor in (2.14).

The only nontrivial component of the curvature form (1.24), for this 2-dimensional theory with the covariant derivative $\partial_j + iqA_j$ is the *magnetic field*

$$B(x) = \partial_1 A_2 - \partial_2 A_1 - ig[A_1, A_2]. \tag{2.26}$$

When we Fourier expand B as in (2.24), its coefficients are given, according to (2.25), by

$$\hat{B}(\vec{p}) = \frac{2}{L}\iint_{\mathbb{T}} \mathrm{Tr}\big(\widehat{\Gamma}(-\vec{p})\partial_1 A_2 - \partial_2 A_1 - ig[A_1, A_2]\big)e^{-i\vec{p}\cdot\vec{x}}d^2\vec{x}. \tag{2.27}$$

Substituting (2.24) the quadratic terms consist of a sum of $\hat{A}_1(\vec{q})\hat{A}_2(\vec{q}\,')$ on $\vec{q}$ and $\vec{q}\,'$ (recall that they are complex numbers and commute) with coefficient

$$\frac{g}{L}\delta(\vec{q}+\vec{q}\,'-\vec{p})F(-\vec{p},\vec{q},\vec{q}\,') \qquad \text{with} \quad F(-\vec{p},\vec{q},\vec{q}\,') = -2i\mathrm{Tr}\Big(\widehat{\Gamma}(-\vec{p})\big[\widehat{\Gamma}(\vec{q}),\widehat{\Gamma}(\vec{q}\,')\big]\Big), \tag{2.28}$$

where we have used the notation of [29].

With the definition (2.14) and the choice of the phase α we have

$$F(\vec{p},\vec{q},-\vec{p}-\vec{q}) = -\sqrt{\frac{2}{N}}\sin\Big(\frac{\overline{n}_{12}NL^2}{4\pi}(\vec{p}\times\vec{q})\Big). \tag{2.29}$$

This is the formula given in [29, (2.29)] and [30, (29)-(30)]. It is convenient for the Hamiltonian approach but for our purpose here it is better to express it in other way. Let $\vec{q} = 2\pi\vec{m}/(LN)$, then, after the relation (2.21), we can substitute the argument of the sine by

$$\frac{\overline{n}_{12}NL^2}{4\pi}\frac{2\pi n_{12}}{LN}(e_2,-e_1)\times\frac{2\pi}{LN}\vec{m}, \tag{2.30}$$

that simplifies to

$$F(\vec{p},\vec{q},-\vec{p}-\vec{q}) = -\sqrt{\frac{2}{N}}\sin\Big(\frac{\pi}{N}(\vec{e}\cdot\vec{m})\Big). \tag{2.31}$$

The important point to keep in mind is that this factor coming from expanding $[A_1, A_2]$ plays the role of the structure constants of the Lie algebra.

Statement of the quantum mechanical problem

Our purpose is to study a pure gauge $SU(N)$ Yang-Mills theory in $\mathbb{T}^2 \times \mathbb{R}$. We have already defined the configuration space $\mathbb{T}^2$, the bundle in terms of a twist n_{12} and we have expanded the field in terms of the electric flux vectors $\vec{e}$ that are representations of the center symmetry, associated to the group $\mathbb{Z}_N \times \mathbb{Z}_N$.

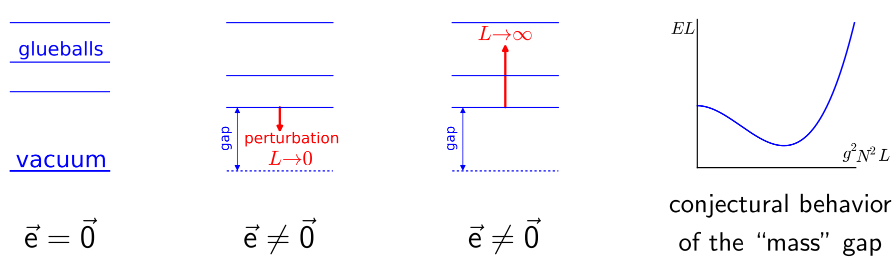
When we consider the quantum mechanical system, the center symmetry is a symmetry of the Hamiltonian that decomposes the underlying Hilbert space $\mathcal{H}$ into a direct sum with the Hamiltonian acting on each *sector*

$$\mathcal{H} = \bigoplus_{\vec{e}} \mathcal{H}_{\vec{e}}, \qquad H_{\vec{e}} \,:\, \mathcal{H}_{\vec{e}} \longrightarrow \mathcal{H}_{\vec{e}}. \tag{2.32}$$

We define as 0 the energy of the ground state in $\mathcal{H}_{\vec{0}}$ that we consider the *vacuum*. The lowest eigenvalues of $\mathcal{H}_{\vec{e}}$ in the rest of the sectors ($\vec{e} \neq \vec{0}$) should remain greater than the vacuum energy, otherwise we have done a wrong choice of the vacuum and we have the *tachyonic instabilities* (the vacuum becomes unstable and decays to a lower state). We shall study these instabilities in the next chapter.

The volume independence in the EK and the TEK models shows a big simplification when $N \to \infty$ in lattice based models, making them independent of the lattice at the limit. The natural question is to study the same phenomenon in a continuous setting. This is the main motivation for the $2+1$ model treated here (in principle $3+1$ is closer to the physical world but it is more difficult). To prevent instabilities, we would like to keep a gap between the vacuum and the spectrum of $H_{\vec{e}}$, $\vec{e} \neq \vec{0}$. This is not possible in full generality (when L and N vary) as the numerical simulations show, but we can play with the twist n_{12} and the conjecture is that for a suitably chosen n_{12} we can keep the spectra above zero for every sector $\vec{e} \neq \vec{0}$ and every L, as N grows (with n_{12} depending on N). Once we have a stable model, volume independence reappears in this context as a dependence on the product NL.

It turns out that for L small the energy of any nonzero sector decreases with L applying perturbation theory but on the other hand, under natural assumptions the energy grows when the volume, L^2, goes to infinity. In a scheme:



In the next section we shall carry out the computations under the perturbative regime but the result will be obtained in a non regularized form that does not show whether it is increasing or decreasing. The regularization and the full study of the tachyonic instabilities is postponed to the next chapter. We cannot fully anticipate here the exact results, we just indicate that the gap is related to

$$x^{-2}|\vec{n}|^2 + \alpha x^{-1} f(\vec{e}/N) + \beta + \gamma x^2 (|\vec{e}|/N)^2 \qquad \text{where} \quad x = \frac{g^2 N^2 L}{4\pi}, \tag{2.33}$$

α, β and γ are essentially constants comparable to 1, and $\vec{n}$ and $\vec{e}$ are related through (2.21). The first two terms come from perturbation theory and the last two terms from the expected non perturbative behavior (confinement). The point is that $f(\vec{v}) \sim -\frac{1}{2}|\vec{v}|^{-1}$ and we have to exploit the arithmetical relation between $\vec{n}$ and $\vec{e}$ to assume the positivity of the gap. This is the novelty of this memoir.

2.3 Perturbation theory

Euclidean regularization of the self-energy

We are going to assume in this section an *Euclidean regularization*, this is a *Wick rotation* in the timelike coordinate: Changing t by $-it$ everywhere, in particular the Minkowski metric $g_{\alpha\beta}$ becomes the Euclidean metric $\delta_{\alpha\beta}$ up to the sign.

The *self-energy* depends on the *vacuum polarization* $\Pi_{\mu\nu}$ and is related to the difference between the propagator and the *dressed propagator*, namely

$$D^{-1}_{\mu\nu}(p) = P^{-1}_{\mu\nu}(p) - \Pi_{\mu\nu}(p). \tag{2.34}$$

This derives from the summation of the *Lippmann-Schwinger series* [54, §7.5]

$$D = P + P\Pi P + P\Pi P\Pi P + \cdots = P(\mathbb{I} - P\pi)^{-1} = (P^{-1} - \Pi)^{-1}. \tag{2.35}$$

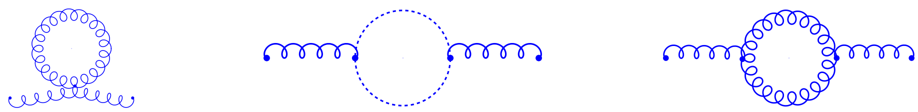
The apparently missing $-i$ factors are due to the Euclidean context we have considered.

In our model there are only massless gluons, represented by the gauge fields A_μ. The gauge fixed Euclidean Lagrangian is

$$\mathcal{L} = \frac{1}{2}\mathrm{Tr}(F_{\mu\nu}F^{\mu\nu}) + \frac{1}{\xi}\mathrm{Tr}(\partial_\mu A_\mu)^2 - 2\mathrm{Tr}(\bar{c}\partial_\mu D^\mu c) \tag{2.36}$$

where ξ is the *gauge parameter* and c is the *ghost field*. If we compare (2.36) to (1.25), the second term is a kind of Lagrange multiplier to impose a particular gauge choice. The last term, the *Faddeev-Popov Lagrangian* has to be still added in the non-Abelian context. It comes from the quantization of the Yang-Mills equations throw the path integral formalism. It is the price to pay to eliminate the redundancy coming from gauge transformations in the non-Abelian case. It is a mathematical artifact, a “ghost” field that violates spin statistics and does not correspond to a physical particle (see [79, III.4] for a specially short and clear explanation).

Each Π in (2.35) corresponds to *one-particle irreducible* [44] and in our case, modeled by (2.36), there are three kind of contributions:

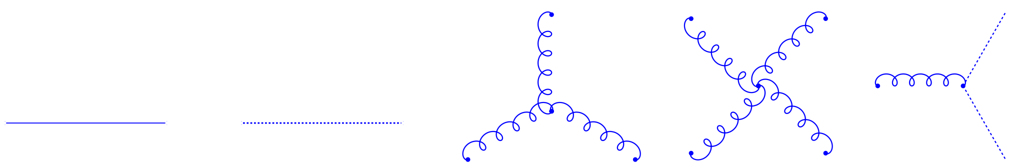


The tadpole The ghost loop The gluon loop

As a matter of fact, the name for the first diagram perhaps is not very appropriate (it is usually applied to the case with one external line).

Some diagrams in QCD and their analogues

The basic *Feynman diagrams* in QCD are



gluon propagator ghost propagator 3-gluon vertex 4-gluon vertex ghost-gluon

See [2] for their contributions. With the usual normalization the *structure constants* f^{abc} are defined by $[T^a, T^b] = if^{abc}T^c$. The orthogonality relation $\mathrm{Tr}\big(T^aT^b\big) = \frac{1}{2}\delta^{ab}$ gives $f^{abc} = -2i\mathrm{Tr}\big([T^a, T^b]T^c\big)$. Comparing this to (2.28), the diagrams for our model can be derived essentially changing the structure constants by F (and $g_{\mu\nu}$ by the Euclidean metric $\delta_{\mu\nu}$).

Following [29], the contribution of the three kind of contributions is

$$\frac{g^2}{L^2}\int_{-\infty}^{\infty}\frac{dq_0}{2\pi}\sum_{\vec{q}} F^2(\vec{p}, \vec{q}, -\vec{p}-\vec{q})\Delta_{\mu\nu}(p, q) \tag{2.37}$$

where for the tadpole and the ghost loop, we have in the Feynman gauge, respectively,

$$\Delta_{\mu\nu}(p, q) = -\frac{2\delta_{\mu\nu}}{q^2} \qquad \text{and} \qquad \Delta_{\mu\nu}(p, q) = -\frac{(p+q)_\nu q_\mu}{(p+q)^2q^2}. \tag{2.38}$$

And for the gluon loop

$$\Delta_{\mu\nu}(p, q) = \frac{\delta_{\rho\rho'}\delta_{\sigma\sigma'}}{(p+q)^2q^2}\widetilde{\Delta}_{\mu\nu}(p, q) \tag{2.39}$$

where $\widetilde{\Delta}_{\mu\nu}(p, q)$ is

(2.40)

$$\big((p+2q)_\mu\delta_{\rho\sigma} + (p-q)_\rho\delta_{\mu\sigma} - (q+2p)_\sigma\delta_{\mu\rho}\big)\big((p+2q)_\nu\delta_{\rho'\sigma'} + (p-q)_{\rho'}\delta_{\nu\sigma'} - (q+2p)_{\sigma'}\delta_{\nu\rho'}\big).$$

The computation of the self-energy

Adding the contribution of the different diagrams [29], the result for the vacuum polarization is

(2.41)
$$\Pi_{\mu\nu} = \frac{g^2}{2L^2}\sum_{\vec{q}}\int \frac{dq_0}{2\pi}\frac{F^2(\vec{p},\vec{q},-\vec{p}-\vec{q})}{q^2(p+q)^2}\big(4(\delta_{\mu\nu}p^2-p_\mu p_\nu)+(p_\mu+2q_\mu)(p_\nu+2q_\nu)-2\delta_{\mu\nu}q^2\big).$$

See [29, (3.48)] and the correction indicated in [30, (27)].

The Euclidean gluon propagator is $P_{\mu\nu}(q)=\delta_{\mu\nu}/q^2$. Then one can read (2.34) as

$$\mathcal{E}^2(\vec{p}) = |\vec{p}|^2 - \sum_\mu \Pi_{\mu\mu} \tag{2.42}$$

with $p=(i\mathcal{E},\vec{p})$ the Euclidean momentum (the usual one after Wick rotation). We define the *self-energy* $\delta\mathcal{E}$ through

$$g^2\delta\mathcal{E}^2(\vec{p}) = -\sum_\mu \Pi_{\mu\mu}. \tag{2.43}$$

According [29, (3.50)], using the *Ward-Takahashi identity* $p_\mu\Pi_{\mu\nu}=0$ and computing the integral in (2.41),

$$g^2\delta\mathcal{E}^2 = \frac{g^2}{2L^2|\vec{p}|^2}\sum_{\vec{q}} F^2(\vec{p},\vec{q},-\vec{p}-\vec{q})\big(2|\vec{p}|^2+\vec{p}\cdot\vec{q}\big)\Big(\frac{1}{|\vec{q}|}-\frac{1}{|\vec{p}+\vec{q}|}\Big). \tag{2.44}$$

We can rearrange the two latter factors as

$$\frac{2|\vec{p}|^2}{|\vec{q}|} - \frac{|\vec{p}|^2}{|\vec{p}+\vec{q}|} + \frac{\vec{p}\cdot\vec{q}}{|\vec{q}|} - \frac{\vec{p}\cdot(\vec{p}+\vec{q})}{|\vec{p}+\vec{q}|}. \tag{2.45}$$

The last two terms vanish after the summation by the odd symmetry $\vec{q}\mapsto-\vec{q}$ and $\vec{p}+\vec{q}\mapsto-(\vec{p}+\vec{q})$. In the second term we apply the translation $\vec{q}\mapsto-\vec{p}+\vec{q}$, that leaves $F^2(\vec{p},\vec{q},-\vec{p}-\vec{q})$ invariant, to get finally

$$g^2\delta\mathcal{E}^2 = \frac{g^2}{2L^2|\vec{p}|^2}\sum_{\vec{q}} F^2(\vec{p},\vec{q},-\vec{p}-\vec{q})\frac{|\vec{p}|^2}{|\vec{q}|}. \tag{2.46}$$

That is, recalling (2.31) and $\vec{q}=2\pi\vec{m}/(LN)$ with $\vec{m}\in\mathbb{Z}^2-\{\vec{0}\}$

$$g^2\delta\mathcal{E}^2 = \frac{g^2}{L}\sum_{\vec{m}\neq\vec{0}}\frac{\sin^2(\pi\vec{m}\cdot\vec{e}/N)}{|\vec{m}|}. \tag{2.47}$$

This is not the end of the story because the series does not converges and requires a regularization. We shall do it in the next chapter (see also Appendix A).

Chapter 3

Tachyonic instabilities

3.1 The source of the instabilities

Regularization of the self-energy

The calculations done in the perturbative regime, see (2.47), prove that the self-energy is essentially given by the formula

$$(3.1)\qquad S(\vec{x}) = \sum_{\vec{m}\neq\vec{0}} \frac{\sin^2(\pi\vec{m}\cdot\vec{x})}{|\vec{m}|} \qquad \text{where} \quad \vec{x} = \frac{\vec{e}}{N}.$$

The series is divergent almost everywhere (it is only convergent for $\vec{x} \in \mathbb{Z}^2$ indeed). There is nothing unnatural about it because self-energy computations are beyond tree level and require some kind of "renormalization".

In [29] and [30] it is applied a dimensional regularization, a zeta regularization, in combination with some properties of the Jacobi θ-functions. Actually the series (3.1) can be expressed in terms of some zeta functions considered by C.L. Siegel in [60, I.§5] generalizing those of M. Lerch and P. Epstein. In particular, it can be proved the meromorphic continuation of

$$(3.2)\qquad \mathcal{Z}(s,\vec{x}) = 2\sum_{\vec{m}\neq\vec{0}} \frac{\sin^2(\pi\vec{m}\cdot\vec{x})}{|\vec{m}|^s}$$

to the whole complex plane with a single pole at $s = 2$.

In Appendix A we give a different (although equivalent) approach that emphasizes the role of some number analytical objects that have been fruitfully employed in other physical topics (see [19]).

We address the reader to the aforementioned papers or Appendix A to get the following results that summarize the regularization of (3.1) through (3.2) and its behavior near the singularities.

- For $\vec{x} \notin \mathbb{Z}^2$, the nonconverging series (3.1) can be defined as the analytic continuation of $\frac{1}{2}\mathcal{Z}(s,\vec{x})$ in (3.2) to $s = 1$.
- Once we have defined (3.1) in this way, if $d(\vec{x})$ is the distance of $\vec{x}$ to the closest point in $\mathbb{Z}^2$, we have

  $$S(\vec{x})d(\vec{x}) \to -1/2 \qquad \text{as} \qquad d(\vec{x}) \to 0. \tag{3.3}$$

  In fact it is possible to prove $S(\vec{x})d(\vec{x}) = -1/2 + O\big(d(\vec{x})\big)$.

Tachyonic energies in the perturbative and non perturbative regimes

Now we are going to study the consequences of this analysis for the computation of the energy.

Recalling (2.42), (2.43) and (2.47), with the notation of (3.1) we have the following formula for the energy (to first order)

$$\mathcal{E}^2(\vec{p}) = |\vec{p}|^2 + g^2\delta\mathcal{E}^2 \qquad \text{with} \quad g^2\delta\mathcal{E}^2 = \frac{\lambda}{2\pi NL} S\big(\frac{\vec{e}}{N}\big) \tag{3.4}$$

where $\lambda = g^2N$ is the 't Hooft coupling (1.48). We define here

$$x = \frac{\lambda NL}{4\pi} \tag{3.5}$$

The momentum is quantized, (2.21), and it can only take values of the form

$$\vec{p} = \frac{2\pi\vec{n}}{LN} \qquad \text{where} \quad \vec{n} \in \mathbb{Z}^2. \tag{3.6}$$

With this notation, the dimensionless version of the energy is

$$\frac{\mathcal{E}^2}{\lambda^2} = \frac{1}{4x^2}|\vec{n}|^2 + \frac{S(\vec{e}/N)}{8\pi^2 x}. \tag{3.7}$$

The important point is that the positivity of $\mathcal{E}^2$ is conditional:

$$\mathcal{E}^2(\vec{p}) > 0 \quad \Leftrightarrow \quad S\big(\frac{\vec{e}}{N}\big) > -\frac{2\pi^2}{x}|\vec{n}|^2. \tag{3.8}$$

Note that the apparent positivity of the formula (3.1), a sum of positive numbers, it is a mirage that it vanishes when we regularize it[1].

Let $d_{\vec{e}}$ be the distance from $\vec{e}/N$ to the nearest lattice point in $\mathbb{Z}^2$. We know by (3.3) that $S(\vec{e}/N) \sim -1/(2d_{\vec{e}})$ for small values of $d_{\vec{e}}$. If $d_{\vec{e}}$ is small enough, then

[1] Perhaps the best known example of this phenomenon is the zeta regularization of 1+2+3+4+... that gives −1/12 and appears in the *Casimir effect* in one dimension [38] [79, I.8].

$-x/d_{\vec{e}}$ could be a large positive number greater than $2\pi^2|\vec{n}|^2$ and the condition (3.8) would not be fulfilled.

In the context of special relativity, a negative square energy corresponds to a *tachyon*, a particle violating causality and traveling faster than light, recall the basic formula $E^2 = (mc^2)^2/(1 - v^2/c^2)$. At some moment in the 20th century it was considered the existence of tachyons as a possibility [20] but the mainstream nowadays is to consider this weird "imaginary energy" as a symptom of some kind of instability in the theory that is dubbed *tachyonic instability*. In our case, if $\mathcal{E}^2(\vec{p}) < 0$ when x is small, then the perturbative theory collapses and we have a problem with the model.

Putting together (3.8) and the previous comments, the existence of tachyonic instabilities in the framework of perturbation theory is unavoidable when

$$0 < |\vec{n}|^2 d_{\vec{e}} \ll 1 \qquad \text{where} \quad d_{\vec{e}} = \operatorname{dist}\big(\frac{\vec{e}}{N}, \mathbb{Z}^2\big). \tag{3.9}$$

Recall that for a fixed n_{12}, $\vec{n}$ and $\vec{e}$ are related by (2.21).

For a complete view of the problem we have to say something about the situation beyond perturbation theory. We are sketchy here, for more detailed information, see [30, §6] and [29, §4].

Even assuming a condition like (3.9) there is a gap in the reasoning, it is not clear the range in which the perturbation scheme applies. Namely, up to what value of x can we trust perturbation theory. For λL large, one expects confinement. In other words, one expects the energy to become proportional to the length

$$\mathcal{E} \sim \sigma L \propto \sigma \frac{x}{\lambda N} \qquad \text{with } \sigma \text{ the string tension.} \tag{3.10}$$

The numerical results support that σ behaves like $\lambda^2 N \phi(|\vec{e}|/N)$ with $\phi(t) = t + O(t^2)$. In a more precise form, the squared non perturbative prediction is

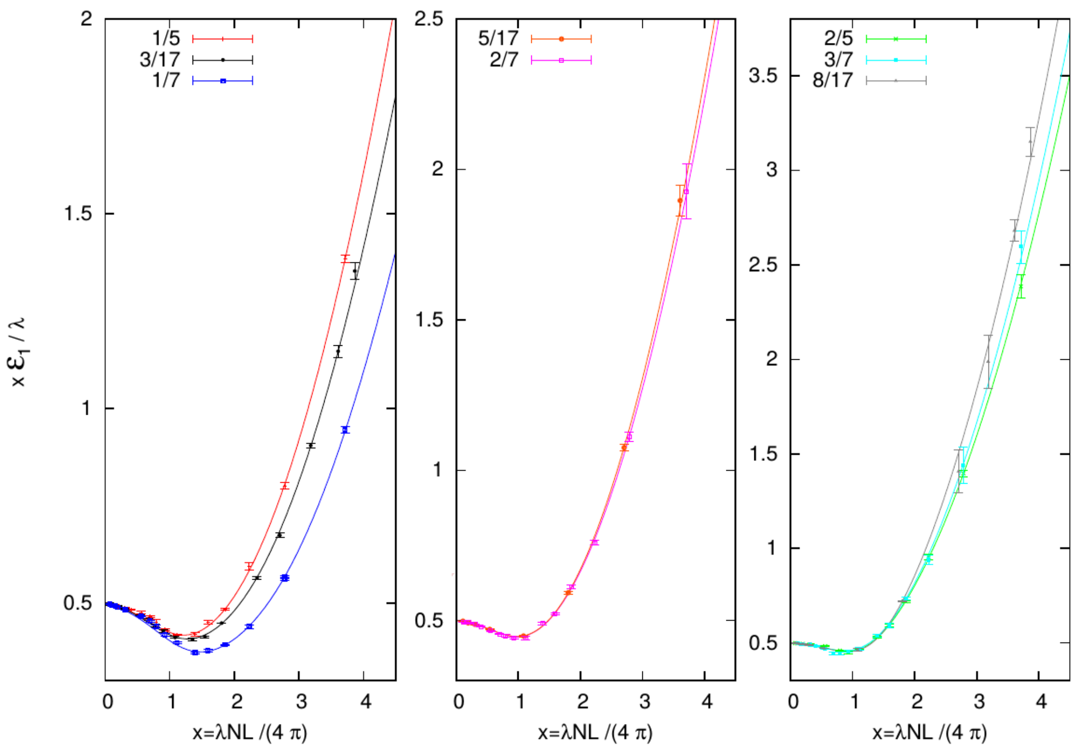
$$\lambda^2 \frac{x^2}{4} \tau^2 \phi^2(|\vec{e}|/N) - \lambda^2 \frac{\tau}{24} \chi(|\vec{e}|/N) \tag{3.11}$$

for certain constant τ, which is dimensionless because λ^2 gives the energy units, and with $\chi(t) = 1 + O(t)$.

A natural way (again supported by the numerical results) of combining the perturbative and the non perturbative regimes in a single formula is simply adding both contributions (3.7), (3.11):

$$\frac{\mathcal{E}^2(\vec{p})}{\lambda^2} = \frac{1}{4x^2}|\vec{n}|^2 + \frac{S(\vec{e}/N)}{8\pi^2 x} + \frac{x^2}{4}\tau^2 \phi^2(|\vec{e}|/N) - \frac{\tau}{24}\chi(|\vec{e}|/N). \tag{3.12}$$

At some small values of N, one can use lattice numerical methods to check this claim. We reproduce here Figure 3 from the preprint of [29]. The plots corresponds to several values of n_{12} and N indicated as a fraction n_{12}/N in the left upper corner



For more on these plots, see [29]

The problem of the tachyonic instabilities arises again in a more general form. If x is not small, the non perturbative terms are relevant. Like in (3.9), we want to find some conditions on $|\vec{n}|$ to detect if (3.12) could be negative for some x. The dominant term for the non perturbative regime is the one with x^2 and we can only play with n_{12} that affects to the choice of $\vec{n}$ in the first term. The worst case scenario (producing tachyonic instabilities) is when the sum of these terms is small and S is negative. Using the inequality of the arithmetic and geometric means, we have

$$\frac{1}{4x^2}|\vec{n}|^2 + \frac{x^2}{4}\tau^2\phi^2(|\vec{e}|/N) \leq \frac{\tau}{2}|\vec{n}|\phi(|\vec{e}|/N) \tag{3.13}$$

and the minimum is attached when both terms in the left hand side are equal. Substituting this bound in (3.12) and the corresponding x where the minimum is attached, we get the following condition for the existence of tachyonic instabilities

$$\frac{\tau}{2}|\vec{n}|\phi(|\vec{e}|/N) + \frac{S(\vec{e}/N)\sqrt{\tau\phi(\vec{e}/N)}}{8\pi^2\sqrt{|\vec{n}|}} - \frac{\tau}{24}\chi(|\vec{e}|/N) < 0. \tag{3.14}$$

Of course, this is only an approximation of the actual condition that would be derived from the minimization of (3.12).

The problematic situation resembles the perturbative case. If $\vec{e}/N$ is close to a point $\mathbb{Z}^2$, the central term takes large negative values and they could be not balanced by the first term. Let us consider that $\vec{e}/N$ is close to $\vec{0}$. Using (3.3), $\phi(t) = t + O(t^2)$, $\chi(t) = 1 + O(t)$, asymptotically the condition becomes

$$\frac{\tau}{2}\Big(\frac{|\vec{n}||\vec{e}|}{N}\Big) - \frac{\sqrt{\tau}}{16\pi^2}\Big(\frac{|\vec{n}||\vec{e}|}{N}\Big)^{-1/2} < \frac{\tau}{24}. \tag{3.15}$$

Then we have instabilities when the term in the parentheses is small enough. The reasoning for $\vec{e}/N$ close to $\vec{0}$ is applicable to any other integral point. Disregarding the constants, the analog of (3.9) for the existence of tachyonic instabilities is

$$0 < |\vec{n}| d_{\vec{e}} \ll 1 \qquad \text{where} \quad d_{\vec{e}} = \operatorname{dist}\big(\frac{\vec{e}}{N}, \mathbb{Z}^2\big). \tag{3.16}$$

Note that (3.9) is stronger than (3.16), as it should be because now we are in a more general framework and then the existence of instabilities is more likely.

Some natural problems in the study of instabilities

Recall from (2.21) that the quantized momentum $\vec{n}$ is a "multiple" of $\vec{e}$. Since $\vec{e}$ comes from the center symmetry $\mathbb{Z}_N \times \mathbb{Z}_N$, we have to understand this multiple modulo N. Of course, for $|\vec{n}|$ smaller, (3.9) and (3.16) are more likely fulfilled. Let us then take

$$\vec{n} = \Big(N\rho\big(\frac{ke_2}{N}\big), -N\rho\big(\frac{ke_1}{N}\big)\Big) \qquad \text{with} \quad \rho(t) = t - \big\lfloor t + \frac{1}{2}\big\rfloor \quad \text{and} \quad k = n_{12}. \tag{3.17}$$

(Here $\lfloor \cdot \rfloor$ is the integral part, as usual). Note that $N\rho(m/N)$ is the residue of m modulo N in the interval $[-N/2, N/2)$ and at the same time $|\rho(t)| = \operatorname{dist}(t, \mathbb{Z})$, the triangular wave.

If we restrict ourselves to $\vec{e} = (e, 0)$ then the conditions (3.9) and (3.16) read

$$\Big|N\rho\big(\frac{ke}{N}\big)\rho\big(\frac{e}{N}\big)\Big| \ll 1 \qquad \text{and} \qquad \Big|\Big(N\rho\big(\frac{ke}{N}\big)\Big)^2\rho\big(\frac{e}{N}\big)\Big| \ll 1. \tag{3.18}$$

Of course, we exclude the case $e = 0$ (equivalently, any zero value modulo N) that corresponds to the vacuum. The assumption $\vec{e} = (e, 0)$ is justified because $\operatorname{dist}(\vec{x}, \mathbb{Z}^2)$ is comparable to $\max\big(|\rho(x_1)|, |\rho(x_2)|\big)$ with $\vec{x} = (x_1, x_2)$, like the L^∞-distance (Chebyshev distance) and the usual distance on the plane.

The problem that we want to study boils down to know if with a suitable choice of $k = n_{12}$ we can avoid the tachyonic instabilities corresponding to (3.18) for all $e \neq 0$. To consider both conditions simultaneously, we introduce

$$a(N, n) = \max_{k\in\mathbb{Z}} \min_{0<e<N} \Big|\Big(N\rho\big(\frac{ke}{N}\big)\Big)^n \rho\big(\frac{e}{N}\big)\Big|. \tag{3.19}$$

The instabilities (in each context) are unavoidable if $a(N,1) \ll 1$ and $a(N,2) \ll 1$. The values $n > 2$ are meaningless in our approach but they could appear in other optimizations (for instance involving higher order terms). The appearance of the sawtooth wave ρ with discontinuities at the integers, gives to the problem of the instabilities a number theoretical twist and we shall benefit from this interplay in §3.2.

There are several problems that can be considered here. The rough general idea is that we would like to avoid instabilities for large values of N, but we could ask for the absence of instabilities for any N or just for a sequence $N_j \to \infty$. In other words, we can study $\lim_N a(N,n)$ or $\limsup_N a(N,n)$. Even if we decide in favor of the second possibility, one has to decide if sequences with very low density, let us say with super-exponential growth, are admitted.

A finer issue is the size of the constants. In (3.9) the symbol $\ll$ hides the range of applicability of perturbation theory and the precise point at which we should consider that the instability happens, depends on the size of the coupling constants.

Taking as granted the absence of instabilities for a particular N, another natural question is to know how much freedom we have in the choice of k. In a stronger form one may ask whether there is a method to compute the value of k at which the maximum is attained. In a weaker and more practical form, it could be interesting to design an algorithm to avoid the exhaustive check of every k modulo N in (3.19). This comment also extends to e if we want a realistic method to compute $a(N,n)$ when N is very large. Naturally the range of k can be restricted to $0 < k < N$ or $0 < k < N/2$ and the same for e (see below), then the trivial algorithm requires $O(N^3)$ steps and it could be unfeasible to run it on a home computer for ranges involving values of N like several tens of thousands. An approximation of $a(N,n)$, or at least some upper and lower bound, are relevant to study the ranges of the possible coupling constants.

To sum up, some of the most natural problems and their physical motivations are in the following list:

- If $a(N,n)$ has a universal lower bound for $n = 1$ as $N \to \infty$, then there are not tachyonic instabilities in the model.
- If a universal lower bound holds for $n = 2$, then the perturbative part is instabilities free.
- The possibility of bounds of this kind for a sequence $N_j \to \infty$ would establish a form a defining a large N limit of the model.
- An algorithm to restrict the possibilities for k and e in (3.19) would be very convenient to carry out numerical studies of the stability of the model.

- An algorithm to find a k in (3.19) perhaps not reaching the maximum but establishing a good lower bound for $a(N,n)$ would give a way of finding for each N a twist n_{12} to avoid instabilities.

3.2 A number theoretical approach

Instabilities and basic Diophantine approximation

Since the rational numbers are dense in $\mathbb{R}$, in principle it seems a naive problem approximating irrationals by rationals or, even more, rationals by rationals, but when the denominators are bounded, there is a plethora of highly nontrivial results and open questions. Let us see firstly that the conditions (3.18) for tachyonic instabilities, lead to problems of this kind through a reformulation of (3.19).

By the definition (3.17) of the function ρ, we have (we use q instead of e for later convenience)

(3.20)
$$\Big|\rho(\frac{q}{N})\Big| = \frac{1}{N}\min_{l\in\mathbb{Z}}|q-lN| \qquad \text{and} \qquad \Big|\Big(N\rho(\frac{kq}{N})\Big)^n\rho(\frac{q}{N})\Big| = \rho(\frac{q}{N})\min_{l\in\mathbb{Z}}|kq-lN|.$$

The latter expression is invariant under

(3.21)
$$(k,l)\mapsto(N-k,q-l) \qquad \text{and} \qquad (q,l)\mapsto(N-q,k-l),$$

then (using that we can assume $0<k<N$), we can restrict ourselves to the case $0<k,q<N/2$. After these considerations, we can rephrase (3.19) as

(3.22)
$$a(N,n) = N^{n-1}\max_{\alpha}\min_{l/q\in\mathcal{F}^*_{N-1}} q^{n+1}\big|\alpha-\frac{l}{q}\big|^n \qquad \text{with} \quad \alpha\in\Big\{\frac{0}{N},\frac{1}{N},\dots,\frac{(N-1)/2}{N}\Big\}$$

where $\mathcal{F}^*_{N-1}$ are the Farey fractions (see below) not exceeding 1/2.

In this way the problems regarding the tachyonic instabilities lead to problems on the approximation by rationals which is the topic of the so-called *Diophantine approximation*. We review here two basic and elementary concepts in this area.

The Farey fractions $\mathcal{F}_M$ are simply the irreducible fractions in $[0,1]$ with denominators less or equal than M. They are usually written as an ordered list called the *Farey sequence*. For instance, for $M=7$

(3.23)
$$\frac{0}{1},\frac{1}{7},\frac{1}{6},\frac{1}{5},\frac{1}{4},\frac{2}{7},\frac{1}{3},\frac{2}{5},\frac{3}{7},\frac{1}{2},\frac{4}{7},\frac{3}{5},\frac{2}{3},\frac{5}{7},\frac{3}{4},\frac{4}{5},\frac{5}{6},\frac{6}{7},\frac{1}{1}.$$

Although it simple definition, the sequence presents some nontrivial properties. The main one, indeed characterizing them, is that the difference of consecutive Farey fractions l/q, l'/q' is $1/qq'$. Combining this with $q+q'>M$, it is possible to deduce [49] that the union of the intervals $\big[l/q-1/qM, l/q+1/qM\big)$ covers $[0,1]$, giving Dirichlet's theorem: For any $\alpha\in\mathbb{R}$ and any $M\in\mathbb{Z}^+$

$$\text{there exists } \frac{l}{q} \text{ with } q<M \text{ such that } \quad \Big|\alpha-\frac{l}{q}\Big|<\frac{1}{qM}. \tag{3.24}$$

The other concept to be reviewed is that of continued fraction. This is a classic topic that appeared in ancient times and played a role in numerical approximations (in part retaken in the 20th century by Padé approximants). Nowadays it is not so widely known beyond number theory.

A *continued fraction* is a (finite or infinite) list of integers, called *partial quotients*,

$$[a_0; a_1, a_2, a_3, \dots] \qquad \text{where } a_0\in\mathbb{Z} \text{ and } a_1, a_2, a_3\dots\in\mathbb{Z}^+, \tag{3.25}$$

meaning the expression $a_0+1/\Big(a_1+1/\big(a_2+(a_3+\dots)\big)\Big)$. The truncated lists

$$\frac{p_n}{q_n}=[a_0; a_1, a_2, a_3, \dots, a_n] \tag{3.26}$$

are called convergents of the continued fractions (assumed irreducible). Very often it is also defined $p_{-1}=1$ and $q_{-1}=0$.

For the numerical calculations it is important to know how to compute efficiently the convergents of the continued fractions from the partial quotients and also the partial quotients from the value of the continued fraction. For the first problem one just notices that p_n and q_n satisfy the recurrence $x_j=a_jx_{j-1}+x_{j-2}$ for $j\in\mathbb{Z}^+$. In some situations, it is convenient to write it in matrix form

$$\begin{pmatrix} a_0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix} a_1 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix} a_2 & 1\\ 1 & 0\end{pmatrix}\cdots\begin{pmatrix} a_j & 1\\ 1 & 0\end{pmatrix}=\begin{pmatrix} p_j & p_{j-1}\\ q_j & q_{j-1}\end{pmatrix} \qquad \text{for } j\ge 0. \tag{3.27}$$

Regarding the second problem, given the value x of the continued fraction, the partial quotients are computed with the following algorithm starting with $x_0=x$, $a_0=\lfloor x\rfloor$

$$x_j=\frac{1}{x_{j-1}-a_{j-1}}, \qquad a_j=\lfloor x_j\rfloor \quad \text{for } j\ge 1. \tag{3.28}$$

The irreducible fractions have finite continued fractions and their partial quotients are actually the quotients in the Euclidean algorithm (this explains the terminology). It is relevant to note that the number of steps of this algorithm is bounded by less than five times the number of digits of the smaller number [49]. Then with appropriate (and easily available) software, computing the continued fraction when the numerator and the denominator have million of digits is a doable task in a home computer.

The relevant properties of the continued fractions for our problem are their approximation properties. In a wide sense, the convergents give the "best approximation" of a real number by rationals. There are two main ways of interpreting the meaning of this assertion, called *first kind approximation* and *second kind approximation* (see [10] for the definitions). Here we focus on the second because it is more convenient for our problem. Let α be a real number and p_j/q_j its convergents, then

$$|q_j\alpha - p_j| = \min_{l/q\in\mathcal{F}_M} |q\alpha - l| \qquad \text{for any} \quad q_j \le M < q_{j+1}. \tag{3.29}$$

It is possible to give an alternative formula for the left hand side [49, §7.5]:

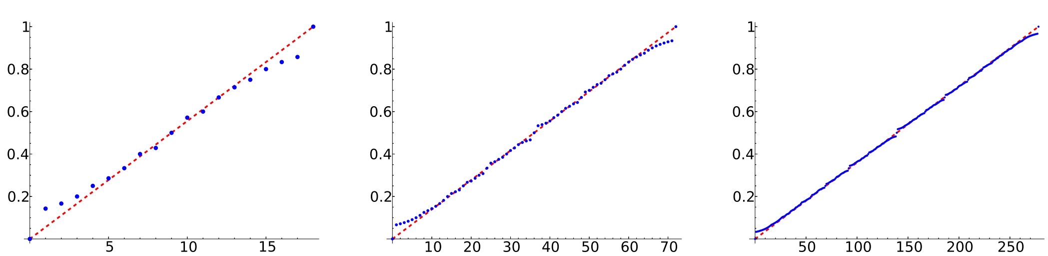
$$q_j\alpha - p_j = \frac{(-1)^j}{\alpha'_{j+1}q_j + q_{j-1}} \qquad \text{with} \quad \alpha'_j = [a_j; a_{j+1}, \dots], \tag{3.30}$$

where a_j are the partial quotients corresponding to α. If α is rational, α'_{j+1} is not defined from certain $j = j_0$ onwards, meaning that $\alpha = p_{j_0}/q_{j_0}$ and there are not more partial quotients to define α'_{j+1}.

Continued fractions and Farey fractions are not unrelated concepts. For instance, taking determinants in (3.27), we have $|p_jq_{j-1} - q_jp_{j-1}| = 1$ that gives the characterizing property of $\mathcal{F}_N$ on the spacing of p_{j-1}/q_{j-1}, p_j/q_j, and then they are consecutive Farey fractions in $\mathcal{F}_{q_j}$. Reciprocally, neighboring Farey fractions have closely related continued fractions.

Comments on the heuristics

Although we have not been able to provide mathematical proofs solving the whole list of problems in the first part of this section, it is clear that the techniques of Diophantine approximation give a deep insight on these problems. We start with some intuitive ideas that can be enough for those less demanding with the rigor and we shall finish with some complete proofs.



$M = 7$, $\#\mathcal{F}_M - 1 = 18$ $\quad$ $M = 15$, $\#\mathcal{F}_M - 1 = 72$ $\quad$ $M = 30$, $\#\mathcal{F}_{30} - 1 = 278$

It is thought that Farey fractions are approximately evenly distributed in $[0, 1]$. For instance, one would expect that $\{k/18\}_{k=0}^{18}$ mimics (3.23), here 18 is the number the pairs of consecutive fractions (the cardinality minus 1). This is not close to truth near the extreme values 0 and 1, but it is approximately correct for many of the points.

As a matter of fact, if d_j is the difference between the j-th of the Farey sequence $\mathcal{F}_M$ and the j-th element of the evenly distributed model, then it is known that

$$\sum_j d_j^2 = O\big(M^r\big) \quad \text{for every } r > -1 \tag{3.31}$$

is equivalent to the Riemann hypothesis (Franel's theorem).

It is well-known [10], the asymptotics

$$\#\mathcal{F}_N \sim \frac{3}{\pi^2}N^2, \qquad \text{in fact} \quad \#\mathcal{F}_N = \frac{3}{\pi^2}N^2 + O(N\log N). \tag{3.32}$$

Then in (3.22) one expects that usually $|\alpha - l/q| \approx q^{-2}$. A weak form of it is already implicit in (3.24) because an elementary calculation proves that the $\alpha \in [0,1]$ that require $q < \epsilon M$ have measure less than 2ϵ

Under the hypothesis that $|\alpha - l/q| \approx q^{-2}$ is the typical situation, for $n > 1$ the minimum should be reached or the greatest denominator, resulting $a(N,n)$ comparable to $N^{n-1}N^{n+1}\big(N^{-2}\big)^n = 1$. This would rule out the possibility of tachyonic instabilities at perturbation level taking $n = 2$ by (3.9). See also the comments after (3.19).

The analysis of the case $n = 1$ is different but with a similar conclusion. According to the stated heuristicsg, $q^{n+1}|\alpha - l/q|^n$ should remain more of less constant, but for large q the statistical fluctuations from the model are expected to be bigger for larger numbers. Then, most of the times the minimum should be reached for small denominators. Although the model still gives $1 = N^0 \cdot 1^2 \cdot 1^{-2}$, preventing the instabilities from happening, this case is not as convincing as the previous one because the model of evenly distributed Farey fractions loses accuracy for small denominators.

The numerical calculations support the model establishing a clear difference between the cases $n = 1$ and $n > 1$ although both numerically appear as non zero lower bounded.

Some new results

For the rigorous partial proofs, we firstly note that (3.29) implies that we can restrict ourselves in (3.22) to l/q belonging to the convergents of $a(N,n)$. This is a major reduction in the computation of $a(N,n)$, fulfilling one of our aims. It lowers the $O(N^3)$ steps of the brute force algorithm to $O(N\log N)$, because there are $O(N)$

possibilities of α to check and for each the Euclidean algorithm to compute the convergents takes $O(\log N)$.

If we denote by $[0; a_1, a_2, \dots, a_M]$ the continued fraction of each valid α, then (3.22) and (3.29) imply

$$a(N,n) = N^{n-1} \max_{\alpha} \min_{j<M} \frac{q_j}{\big(\alpha'_{j+1} q_j + q_{j-1}\big)^n} \qquad \text{with} \quad \alpha'_j = [a_j; a_{j+1}, \dots, a_M]. \tag{3.33}$$

We consider the case $n = 1$. The upper bounds are not so relevant for the problem of the instabilities but the nature of the optimal ones admits a neat solution that is worthy to mention.

Recall the *Fibonacci sequence* $\{F_k\}_{k=1}^{\infty} = (1, 1, 2, 3, 5, 8, \dots)$ defined by the recurrence $F_{j+2} = F_{j+1} + F_j$. Assume that $N \geq 5$ is a Fibonacci number, say $N = F_J$. Thanks to (3.27), it is easy to see that

$$[0; 2, 1, \overset{j-5 \text{ times}}{\dots\dots\dots}, 1, 2] = \frac{F_{J-2}}{F_J} = \frac{F_{J-2}}{N}. \tag{3.34}$$

Let us call this number α_0. We are going to check that $a(N, 1) = \alpha_0$. The convergents are $0/1, 1/2, 1/3, 2/5, \dots, F_{J-4}/F_{J-2}$ and α_0. In the same way, $\alpha'_1 = 1/\alpha_0$, $\alpha'_j = F_{J-j}/F_{J-j-1}$. Then for $\alpha = \alpha_0$ the minimum is reached for $j = 0$ giving α_0. By (3.33), to deduce $a(N, 1) = \alpha_0$, it remains to prove that for any $\alpha \neq \alpha_0$ (with denominator N) there exists j_0 such that

$$q_{j_0} \leq \big(\alpha'_{j_0+1} q_{j_0} + q_{j_0-1}\big)\alpha_0. \tag{3.35}$$

If α has a partial quotient greater that 2, this holds true because $\alpha_0 > 1/3$ and we can take $\alpha'_{j_0+1} = [3; \dots] \geq 3$. Otherwise, the partial quotients of the α'_j are 1 or 2. If $\alpha \neq \alpha_0$, there exists an α'_{j_0+1} of the form $[2; 1, \dots]$, $[2; 2, \dots]$ or $[2; 2] = [2; 1, 1]$. In any of these cases $\alpha'_{m_0} > 7/3$ and (3.35) is fulfilled.

Revising the proof, one notes that the assumption $N = F_J$ was only employed to compute the value at $\alpha = \alpha_0$. The proof still applies for $N > F_J$ except that $\alpha = \alpha_0$ is not a valid value in (3.22). Then we have proved

$$a(N, 1) \leq \frac{F_{J-2}}{F_J} \quad \text{if} \quad N \geq F_J, \quad \text{with equality if } N = F_J. \tag{3.36}$$

The last part leads to ask if there are infinitely many Fibonacci prime numbers. This is an open question considered very hard (an exponential problem) in number theory. The well-known asymptotic

$$F_j \sim r^j \quad \text{as } j \to \infty \quad \text{with } r = \frac{1+\sqrt{5}}{2} \quad \text{(the golden ratio)}, \tag{3.37}$$

gives the asymptotic bound $a(N,1) \le r^{-2} = 0.381966\dots$, $N \to \infty$. The values reaching the bound (3.36) appear in the numerical data as outliers because of the exponential growth in (3.37).

For the lower bound, we appeal to *Zaremba's conjecture*, this is a problem posed in 1971 (see an overview in [43]) that remains open yet. It claims the existence of a positive integer A with the following property:

(3.38)
$$\text{for every } N \in \mathbb{Z}^+, \text{there exist } a_1, a_2, \dots, a_j \le A \text{ such that } [0; a_1, a_2, \dots, a_j] = \frac{p_j}{N}.$$

In a more elementary way, it means that the recurrence $q_{j+1} = a_{j+1}q_j + q_{j-1}$, $q_0 = 0$, $q_1 = 1$ can capture any positive integer with a judicious choice of the a_j.

Recently there was a breakthrough on (3.38). In [6] it has been proved that $A = 50$ is valid for any N except for a zero density set[2]. A common ansatz in number theory is that the prime numbers are random and they behave like the integers for properties not involving the multiplicative structure. If it is applicable here, we could conclude the absence of tachyonic instabilities in the model for almost every N from [6]. Indeed the lower bound for $a(N,1)$ is a straightforward consequence of the formula (3.33) under (3.38), because

$$a(N,1) > \max_{\alpha} \min_{j} \frac{1}{\alpha'_{j+1} + 1} \ge \frac{1}{A+2}. \tag{3.39}$$

A kind of converse is also true: using $\alpha'_{j+1} \ge a_j$, if $a(N,1) > \epsilon$ then we could take $A = \lfloor \epsilon^{-1} \rfloor$ in (3.38) for that N. It is thought that $A = 2$ for a certain (large value of) N onwards. If this is true, one argument that we do not reproduce here here (essentially bounding $q_j/q_{j-1} > 5/4$) would lead to $a(N,1) > 5/19$ for large N.

Now we treat the case $n = 2$. We are going to prove

$$a(N,2) > \frac{3}{\pi^2} = 0.30396355\dots \tag{3.40}$$

The method below gives a slightly better bound. Our interest here is just proving the absence of the instabilities in the perturbative regime in the form stated in (3.9).

For each k, let

$$F(k) = \min_{l/q} q^3 \left| \frac{k}{N} - \frac{l}{q} \right|, \qquad \text{hence} \quad a(N,2) = \max_{0<k<N} F(k). \tag{3.41}$$

Given k, say that the minimum in (3.41) is attained at l_k/q_k. We define the sets

$$\mathcal{C}_m = \big\{ 0 < k < N \ : \ q_k k - l_k N = m \big\} \qquad \text{with} \quad 0 \ne |m| < N/2. \tag{3.42}$$

[2]If there are E_N exceptions less than N, then $E_N/N \to 0$ as $N \to \infty$.

Note that, coming back to the original formulation of the problem (3.18), we have $m = N\rho(ke/N)$ with $e = q_k$. Clearly, in this range of m,

$$\sum_m \#\mathcal{C}_m \geq \#\Big(\bigcup_m \mathcal{C}_m\Big) \geq \#\{0 < k < N\} = N - 1. \tag{3.43}$$

On the other hand, if $k \in \mathcal{C}_m$ then $F(k) = N^{-1} q_k |m|^2$, consequently

$$q_k \leq \frac{N}{m^2} \max_{k\in\mathcal{C}_m} F(k) \leq \frac{N}{m^2} a(N,2) \quad \text{that implies} \quad \#\big\{q_k \ : \ k \in \mathcal{C}_m\big\} \leq \Big\lfloor \frac{N}{m^2} a(N,2) \Big\rfloor. \tag{3.44}$$

But we also have $q_k \equiv m\bar{k} \pmod N$, then $\big\{q_k \ : \ k \in \mathcal{C}_m\big\}$ has the same cardinality as $\mathcal{C}_m$ (recall $0 < q_k < N$). From these observations and (3.43), we get

$$N-1 \leq \sum_{0\neq|m|<N/2} \Big\lfloor \frac{N}{m^2} a(N,2) \Big\rfloor \leq 2\sum_{m=1}^{\infty} \frac{N}{m^2} a(N,2) - 2 \sum_{m>\sqrt{Na(N,2)}} \frac{N}{m^2} a(N,2). \tag{3.45}$$

The tail series can be lower bounded by 1, just approximating by the integral. Then

$$N - 1 \leq 2\zeta(2) N a(N,2) - 2 \tag{3.46}$$

and the evaluation of $\zeta(2)$ gives (3.40).

A similar analysis in the case $n > 2$ without taking care of the tail series, produces

$$N - 1 \leq \sum_{0\neq|m|<N/2} \Big\lfloor \frac{N}{m^n} a(N,2) \Big\rfloor \leq 2\zeta(n) a(N,n). \tag{3.47}$$

On the other hand, as mentioned before, the last convergent of α different from itself and α are consecutive Farey fractions. If we take it as l/q in (3.22), we have $|\alpha - l/q| = (qN)^{-1}$ and $a(N,n) < 1/2$ because $q < N/2$. In this way, we have the upper and lower bounds

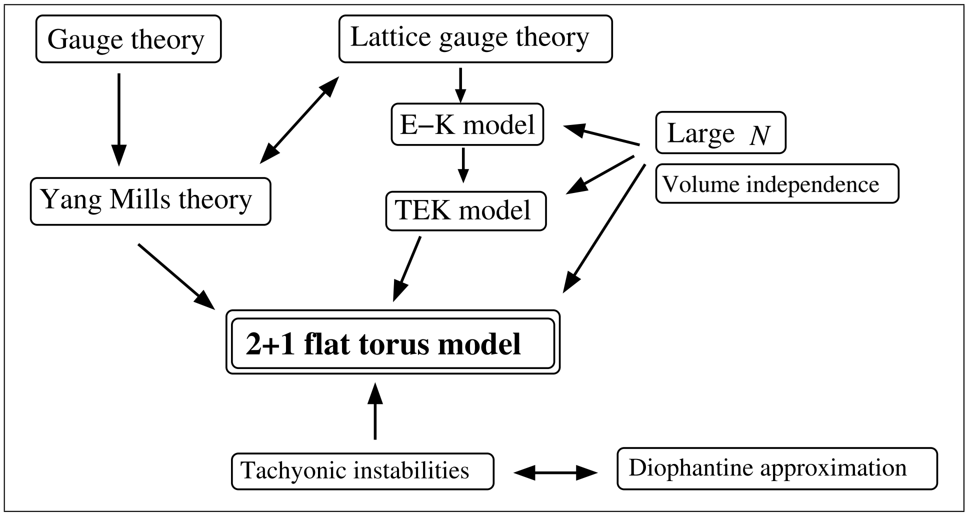
$$\frac{1}{2\zeta(n)}\big(1 - N^{-1}\big) < a(N,n) < \frac{1}{2} \qquad \text{for } n > 1. \tag{3.48}$$

If N and n go to infinity, $a(N,n)$ tends to be $1/2$. The first values of the leading constant in the left hand side are

n	$1/(2\zeta(n))$
2	0.303963551
3	0.415953686
4	0.461969201
5	0.482193670
6	0.491476296

3.3 Summary and conclusions

We start including a scheme summarizing the topics treated in this work and their relations.



Gauge theory is the modern paradigm to study interactions in particle physics. In this context, Yang-Mills theories have become central topics in theoretical physics. They cover Abelian and non-Abelian generalizations of electromagnetism and it is possible to carry out non perturbative numerical calculations with them through discretizations known under the general name of *lattice gauge theory*. For the gauge group $SU(N)$, the limit when the number of colors N goes to infinity is considered a previous step to understand the theories with finite N. It approximates the *meson* phenomenology and also reflects some expected properties of *glueballs* and *baryons* [14]. It also seems to have some vague resemblances to some aspects of string theory [1].

The *Eguchi-Kawai model* (E-K) introduced by T. Eguchi and H. Kawai involves large N expansions and lattice gauge theory. Its main feature is the *volume independence* (reduction): In the large N limit the lattice becomes irrelevant and the configuration space can be reduced to a point. But the model was shown to be defective [4] because the symmetry is spontaneously broken before reaching the limit. To solve this issue, a new model, the *twisted Eguchi-Kawai model* (TEK), was introduced shortly after [34] [33] by A. González-Arroyo and M. Okawa. The important feature to prevent symmetry breaking (at least in some situations [71], see extensive numerical calculations in [36]) was a twist in the periodic boundary conditions. It has antecedents in the paper [66] by G. 't Hooft. In joint works with M. García Pérez [29] [30], the authors of the twisted Eguchi-Kawai model study volume independence for the large N limit of Yang-Mills theories on twisted flat tori without essential reference to lattice gauge theory. The $2 + 1$ dimensional case is treated with special care. This is the model considered in this work.

The underlying Hilbert space in the 2+1 flat torus model is decomposed into independent sectors through electric flux 2-vectors (representing the *center symmetry*) whose components are integers modulo N. There are also topological sectors, *twist sectors*, in the space of bundles with gauge group $SU(N)/Z_N$, where Z_N is the center of $SU(N)$ (cf. [66]). They are characterized by an integer modulo $N$and it can be interpreted as a discrete flux traversing the torus. To avoid the existence of subgroups, one assumes N to be a prime number. One can employ the usual methods (Feynman diagrams) in the perturbative regime while in the non perturbative regime one can appeal to dimensional analysis and the expected confinement. There is also an expected way of matching both regimes. It is not clear if the model has *tachyonic instabilities* (imaginary energies) neither in the perturbative regime nor in general. The question we address in this work is if, for N large, the model is tachyonic instability free in one of the twist sectors. The existence of these instabilities would represent a kind of phase transition for large N that would prevent volume independence.

The conclusions of our study are as follows:

- After some reductions, the problem becomes a number theoretical problem. Essentially, in a weighted form, how far apart can the fractions of denominator N be from the fractions with smaller denominator. It appears in two flavors, one to avoid tachyonic instabilities in the perturbative regime and another stronger one, to avoid them in general.

- We show that in this interplay, the absence of tachyonic instabilities or any N is equivalent to an open problem in number theory (Zaremba's conjecture).

- It is also shown, that when N is a Fibonacci prime numbers then an optimal bound is reached. This is the best situation to avoid instabilities.

- The numerical calculation of the expression characterizing the existence of instability is greatly reduced using continued fractions.

- We solve the weak form of the problem, that implies that the model is free of tachyonic instabilities in the perturbative regime in some twist sector.

We are composing a research joint paper [9] covering the above conclusions.

Appendix A

A simple approach to regularization

We include here an approach to the regularization of (3.1). We do not claim any advantage of this method with respect to [29]. It just illustrates how can be obtained in a more or less elementary or intuitive way (without appealing to Jacobi theta functions) combining some ideas in physics, analysis and number theory.

We can write $\mathcal{Z}(s,\vec{x})$ in (3.2) as

$$\text{(A.1)} \qquad \mathcal{Z}(s,\vec{x}) = Z(s,\vec{0}) - Z(s,\vec{x}) \qquad \text{with} \quad Z(s,\vec{x}) = \sum_{\vec{k}\neq\vec{0}} \frac{e^{2\pi i\vec{k}\cdot\vec{x}}}{|\vec{k}|^s}$$

where we have simply written $\sin^2\alpha = \frac{1}{2}\Re(1-e^{i\alpha})$ and the convergence is assured for $\Re(s) > 2$. The series $Z(s,\vec{0})$ is the simplest instance of Epstein zeta function and it is known that it admits a meromorphic continuation to the whole $\mathbb{C}$ plane with a simple pole at $s = 2$. Instead to appealing to this result [59], we proceed using elementary arguments that can be traced back to C.F. Gauss.

Consider $r_2(n) = \#\{(a,b) \in \mathbb{Z}^2 \ : \ a^2+b^2 = n\}$, the number of representations as a sum of two squares, then

$$\text{(A.2)} \qquad Z(s,\vec{0}) = \sum_{n=1}^{\infty} \frac{r_2(n)}{n^{s/2}} = \sum_{n=1}^{\infty} \frac{r_2(n)-\pi}{n^{s/2}} + \pi\zeta(s/2),$$

where, as usual, ζ is the Riemann zeta function, the analytic extension of $\sum n^{-s}$. Note that $\sum_{n=1}^{N}\big(r_2(n)-\pi\big) = \#(C\cap\mathbb{Z}^2) - |C|$ where C is a circle of radius $\sqrt{N}$ and $|C|$ is its area. It is clear that this difference should be small. The rather obvious bound $O(\sqrt{N})$ was noticed by Gauss. With very ingenious geometrical-arithmetical (but still elementary) arguments [40] one can get Sierpiński's bound $O\big(N^{1/3}\big)$. The optimal bound is the content of the Gauss circle problem that is still open and under active research. Using Abel summation[1] and Sierpiński's bound in (A.2), we have that $Z(s,\vec{0})$ admits a meromorphic continuation to $\Re(s) > 2/3$ with a simple pole at $s = 2$ (coming from ζ) with residue 2π.

If one of the coordinates of $\vec{x}$ is an integer, then $Z(1,\vec{x})$ strictly speaking does not converge as an iterated sum because the sum on the corresponding coordinate of $\vec{k}$ is essentially the harmonic series. In the rest of the cases, by Abel summation in both variables, we get a converging result because $S(u,v) = \sum_{k_1\le u} e(k_1x_1)\sum_{k_2\le v} e(k_2x_2)$ is bounded. The form of (3.1) suggests a spherical summation to achieve the convergence. If we group the $\vec{k}$ according to the value of $n = k_1^2 + k_2^2$, we have

(A.3)

$$Z(s,\vec{x}) = \sum_{n=1}^{\infty}\frac{1}{n^{s/2}}\sum_{k_1^2+k_2^2=n} e^{2\pi i\vec{k}\cdot\vec{x}} = \frac{s}{2}\int_1^\infty \frac{D(u)}{u^{s/2+1}}\,du \quad \text{with} \quad D(u) = \sum_{k_1^2+k_2^2\le u} e^{2\pi i\vec{k}\cdot\vec{x}}.$$

In first approximation (use a smoothing and the Poisson summation formula [12] for a rigorous treatment), for $|\vec{x}|$ small we have

$$D(u) \sim \iint_{|\vec{y}|^2\le u} e^{2\pi i\vec{y}\cdot\vec{x}}\,d\vec{y} = \frac{\sqrt{u}}{|\vec{x}|}J_1(2\pi|\vec{x}|\sqrt{u}). \tag{A.4}$$

This is a well-known calculation in Fraunhofer diffraction [5] (in fact, stretching the analogy, $D(u)$ is a lattice approximation to Huygens principle). Then we have

$$Z(s,\vec{x}) \sim \frac{s}{2|\vec{x}|}\int_1^\infty u^{-(s+1)/2}J_1(2\pi|\vec{x}|\sqrt{u})\,du = s|\vec{x}|^{s-2}\int_{|\vec{x}|}^\infty t^{-s}J_1(2\pi t)\,dt. \tag{A.5}$$

The last integral converges for $\Re(s) > 0$ and uniformly in $|\vec{x}|$ for any $0 < \Re(s) < 2$. For $s = 1$ we have

(A.6)

$$Z(1,\vec{x}) \sim |\vec{x}|^{-1} \quad \text{as } |\vec{x}|\to 0, \quad \text{in general,} \quad Z(1,\vec{x}) \sim |\vec{x}-\vec{x}_0|^{-1} \quad \text{as } \vec{x}\to\vec{x}_0\in\mathbb{Z}^2$$

because $\int_0^\infty t^{-1}J_1(2\pi t)\,dt = 1$ [39, 6.623]. The important point here is not the value of the constant but the singular behavior for $s = 1$. If we are not interested on the constant, the singularity can be guessed from a less rigorous but completely elementary argument. If instead of (A.4) we use a "top-hat" approximation saying that for $|\vec{x}| < C/\sqrt{u}$, $D(u)$ is essentially the area of the circle defined by the limits of the summation, πu, and it is negligible otherwise due to wave interference, then, keeping in mind (A.3), one hopes

$$Z(1,\vec{x}) \sim \frac{1}{2}\int_1^\infty \frac{D(u)}{u^{3/2}}\,du \propto \int_1^{C^2/|\vec{x}|^2}\frac{\pi u}{u^{3/2}}\,du \propto |\vec{x}|^{-1} \tag{A.7}$$

[1] We mean the elementary identity $\sum_{n=1}^N a_nf(n) = A(N)f(N) - \int_1^N A(x)f'(x)\,dx$ with $A(x) = \sum_{n\le x} a_n$. If $A(N)f(N)\to 0$ one gets $\sum_{n=1}^\infty a_nf(n) = -\int_1^\infty A(x)f'(x)\,dx$.

that matches (A.6).

As a subproduct of our analysis, one gets the formula

$$\text{(A.8)} \qquad S(\vec{x}) = K - \frac{1}{4}\int_1^\infty \frac{D(u)}{u^{3/2}}\,du \qquad \text{with} \quad K = \frac{1}{2}\sum_{n=1}^\infty \frac{r_2(n)-\pi}{n^{1/2}} + \frac{\pi}{2}\zeta(1/2)$$

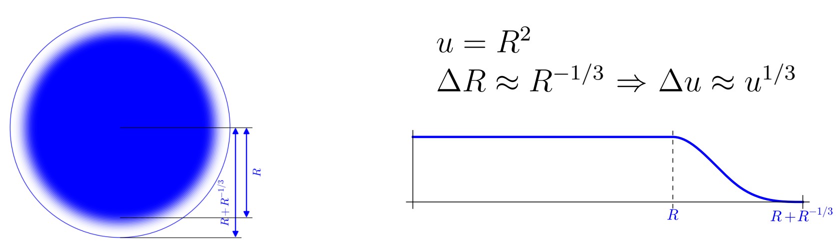
and $D(u) = \sum_{|\vec{k}|^2\le u} e^{2\pi i \vec{k}\cdot\vec{x}}$. It is possible to give an "explicit" formula for K if one admits Dirichlet L-functions modulo 4, $L(s)$, namely[2] $K = 2\zeta(1/2)L(1/2)$. A noteworthy point is that the approximation to $S(\vec{x})$ around $\vec{x} = \vec{0}$ is radial but $S(\vec{x})$ it is not. In Appendix B we illustrate graphically the situation.

Once we know that (A.8) is meaningful, it becomes quite explicit. Note that the integral can be calculated in each interval $[n, n+1)$ with n integer (because D is constant there). In this way, $S(\vec{x})$ is expressed as an oscillatory sum and its rate of convergence can be improved at will using Fourier series acceleration techniques. We do not claim any numerical advantage with respect to the theta expression given in (3.62) of [29], we just point out the simplicity of the expression.

A loose end in the previous argument is the convergence in (A.8). Note that (A.4), (A.7) are based on a not totally justified cancellation that allows to avoid large values of u.

To assure the convergence in (A.8) it would be enough $D(u) = O(u^\alpha)$ with $\alpha < 1/2$ (assuming $\vec{x} \notin \mathbb{Z}^2$). We give a brief argument to justify $\alpha = 1/3$. It gives some idea about the convergence rate in (A.8). Elaborating this argument, one gets a good behavior of the accelerated series that become suitable for numerical calculations.

Imagine that we blur the circle of summation of radius $R = \sqrt{u}$ in a thin corona of width $R^{-1/3}$ to get a smooth profile.



With this approximation we have $D(u) = D_s(u) + O(u^{1/3})$ with D_s the smoothed sum. Replace $\vec{x}$ (that it is assumed fixed and not in $\mathbb{Z}^2$) by $\vec{x}+\vec{t}$ with $\vec{t}$ a variable. Expanding D_s into Fourier series and substituting $\vec{t} = \vec{0}$, we have that D_s is a sum of Fourier coefficients $\sum_{n,m} a_{nm}$. By the uncertainty principle, the coefficients with $\sqrt{n^2+m^2} = O\big(R^{1/3}\big)$ (that do not see the "details" of size less than $R^{-1/3}$) are

[2] Here $L(1/2)$ is simply the converging series $1^{-1/2} - 3^{-1/2} + 5^{-1/2} - 7^{-1/2} + \dots$

essentially the same as for the non smoothed case, that by (A.4) decays as[3] $u^{1/4}(n^2 + m^2)^{-3/4}$ and the terms with $\sqrt{n^2+m^2} \gg R^{1/3}$ are negligible by the smoothness (integration by parts). In this way, $D(u)$ is bounded by $u^{1/4}\sum_{nm}(n^2+m^2)^{-3/4}$ with $n^2+m^2 < u^{1/3}$ that is $O(u^{1/3})$.

[3]We are using that $Rx^{-1}J_1(Rx) < CR^{1/2}x^{-3/2}$.

Appendix B

Numerical calculations

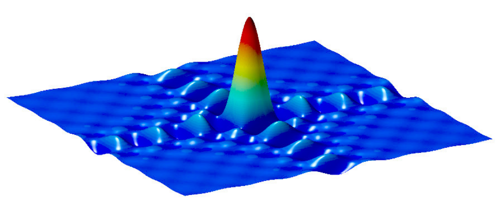
We devote this appendix to develop some topics appearing in this thesis, mainly related to numerical calculations.

Rectangular and circular summation. In the regularization of the self-energy, circular summation played an important role through the function introduced in the formula (A.3)

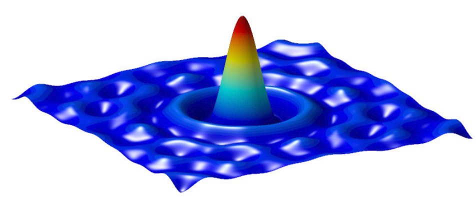
$$D(u) = \sum_{|\vec{k}|^2 \le u} e^{2\pi i \vec{k}\cdot\vec{x}}. \tag{B.1}$$

There are several issues to be considered here. The most practical one is the computation of $D(u)$ and the most theoretical, treated latter, is relative to the convergence in (A.8).

If we fix u and consider D as a function of $\vec{x}$, the behavior around the origin was important for the analysis in §3.2 of the tachyonic instabilities. On the other hand, in §3.1 we mentioned that iterated summation was not enough for the regularization of the self-energy, then there is something in D that behaves better that the same summation on rectangles and still gives the adequate representation of the singularity. This is related to the curvature of the domain but it would be lengthy to enter into details in this point.



Summation over $|k_1| \le 10$, $|k_2| \le 10$



Summation over $k_1^2 + k_2^2 \le 10^2$

We simply exemplify the situation with two graphs of $f(\vec{x}) = \sum_{\vec{k}} e^{2\pi i \vec{k}\cdot\vec{x}}$. The first one with the summation restricted to a square and the second one to a circle. Note the alignment of the bumps in the first case.

Plotting these kind graphs leads to the numerical calculation of D. In principle there are asymptotically πu terms involved in the summation and one may suspect that a number of operation as $O(u)$ is optimal, but it can be lowered to $O(\sqrt{u})$. The point is using the evaluation of the Dirichlet kernel (just computing the sum of a geometric progression)

$$\sum_{k=-N}^{N} e^{2\pi ikx} = \frac{\sin\big(\pi(2N+1)x\big)}{\sin(\pi x)}. \tag{B.2}$$

If we call $\mathcal{D}(a,x)$ the value of this expression with $N = \lfloor a \rfloor$, then

$$D(u) = \mathcal{D}\big(\sqrt{u}, y\big) + 2\sum_{1\le k_1\le\sqrt{u}} \cos(2\pi k_1 x)\mathcal{D}\Big(\sqrt{u-k_1^2}, y\Big) \tag{B.3}$$

where we have employed the formula $e^{it} + e^{-it} = 2\cos t$ to reduce the range of the summation.

It can be implemented in a simple C function as:

```c
double exp_sum(double x, double y, int u){
        double s;
        int k1, j;
        int sqr = (int)sqrt(u);
        s = sin( M_PI*(2*sqr+1)*y )/sin(M_PI*y);

        for(k1=1; k1< sqr+1; ++k1){
                j = (int)sqrt(u-k1*k1);
                s += 2.0*cos(2.0*M_PI*x*k1) * sin( M_PI*(2*j+1)*y )/sin(M_PI*y);
        }
        return s;
}
```

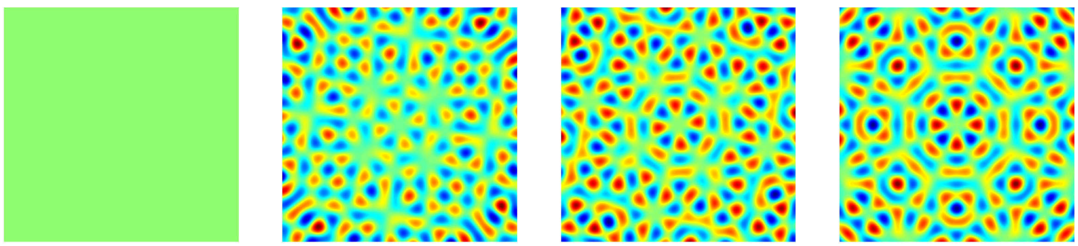
As a matter of fact, the plots above were done storing the results generated by a C program in a file (`bumpdat` a matrix containing the data in a grid) and running the following Matlab/Octave program[1]

```matlab
% load data
bumpdat;
N = size(B,1);
[X,Y] = meshgrid (linspace(-0.5,0.5, N),linspace(-0.5,0.5, N));
Z = sin( pi*(2*R+1)*X).*sin( pi*(2*R+1)*Y)./sin(pi*X)./sin(pi*Y);
% plot
figure(1)
surf(linspace(-0.5,0.5, N),linspace(-0.5,0.5, N),B','EdgeColor','none')
colormap jet
camlight left; lighting phong
axis off
% plot
figure(2)
surf(X,Y,Z,'EdgeColor','none')
colormap jet
camlight left; lighting phong
axis off
```

[1] The commands `camlight left` and `lighting phong` apparently are not still implemented in the current version of Octave.

Radial versus 8-fold symmetry. The function (3.1), regularized in (A.8), is not radial but when $\vec{x}$ is close to a point of $\mathbb{Z}^2$ we are approximating it by a radial function around that point. This may sound strange because there is nothing in the formula (3.1) suggesting a symmetry beyond the obvious ones $x \mapsto \pm x$, $y \mapsto \pm y$, $x \leftrightarrow y$ that are the rigid motions leaving the unit square invariant (the dihedral group D_4).

To get some intuition, in the following figures we show a contour plot (a colored zenith view) on the domain $[-1/2, 1/2] \times [-1/2, 1/2]$ of the function $f\big(r_\theta(\vec{x})\big) - f(\vec{x})$ with $f(\vec{x}) = \sum_{|\vec{k}|<10} e^{2\pi i \vec{k}\cdot\vec{x}}$, i.e. $D(10)$ as a function of $\vec{x}$, where $r_\theta(\vec{x})$ means the result of rotating $\vec{x}$ under an angle θ.



$\theta = 0,\ \pi/2$ $\quad$ $\theta = \pi/12,\ 5\pi/12$ $\quad$ $\theta = \pi/6,\ \pi/3$ $\quad$ $\theta = \pi/4$

In the first figure we have, of course, the background color corresponding to the zero level. The function f has the natural 8-fold symmetry corresponding to the invariance of the original function under D_4, that it is inherited for that of $\mathbb{Z}^2$. The point to be noted here is that regardless the angle, we have a central part surrounded by 8 slices with the background color. It means that in the circle bounded by the slices, f is in practice invariant by rotations in opposition to the kaleidoscopic behavior in the rest of the points. According to the heuristics in (A.7) the radius of this circle is comparable to the inverse of the maximal radius of $\vec{k}$.

The data for the plots were generated using the C function that we mentioned before, evaluated at rotated points with a loop involving

```
exp_sum(cos(theta)*x-y*sin(theta),sin(theta)*x+y*cos(theta),R);
```

The results are stored into matrices and the actual plots were drawn with simple Matlab/Octave lines like

```
surf(linspace(-0.5,0.5, N),linspace(-0.5,0.5, N),(B2-B0)','EdgeColor','none')
colormap jet; view([180 90]); axis square; axis off
```

The numerical calculation of $a(N, n)$. We have already seen that in (3.22) we can also take l/q as a convergent of α by (3.29). The computation of the convergents is as simple as the computation of the greatest common divisor. One can write easily high speed code but for single calculations any mathematical package probably will get the same performance.

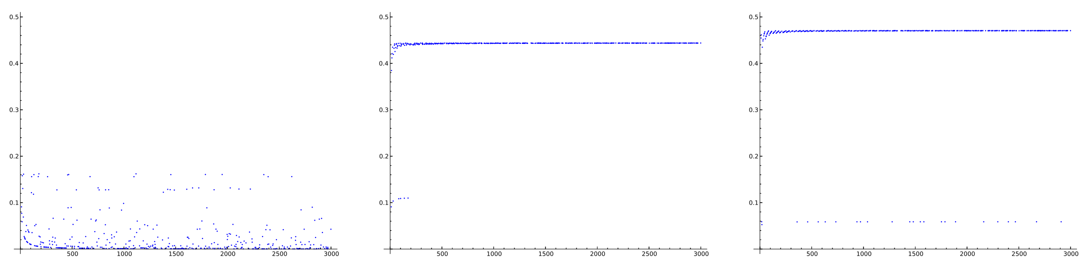
For instance, with the open-source and freely available mathematics software system `SageMath`, one can find with the following lines the values of $k = \alpha N$, l and q (list variable `klpm`) at which the maximum and the minimum are reached in (3.22).

With a simple program like this we can compute $a(N, n)$ and check that the heuristics claimed above is experimentally true, showing a different regime for $n = 1$ and $n > 1$.

```
rmax = 0
for k in srange(1,(N+1)/2):
        r = N^2
        L = (continued_fraction(k/N)).convergents()
        L[:]= L[:-1]

        for frac in L:
                v = frac.denominator()*(abs(frac.denominator()*k/N-frac.numerator()))^n
                if v < r:
                        r = v
                        klp = (k,frac.numerator(),frac.denominator())
        if r>rmax:
                rmax = r
                klpm = klp
```

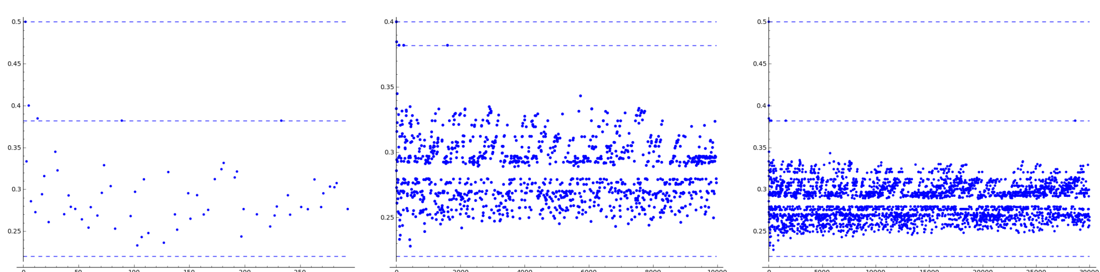
For instance, if we plot k/N in the cases $n = 1$, $n = 2$ and $n = 3$



n = 1 n = 2 n = 3

It reflects the heuristics that for $n > 1$ the minimum is reached typically by the last but one convergent and for $n = 1$ the small denominator convergents are more relevant.

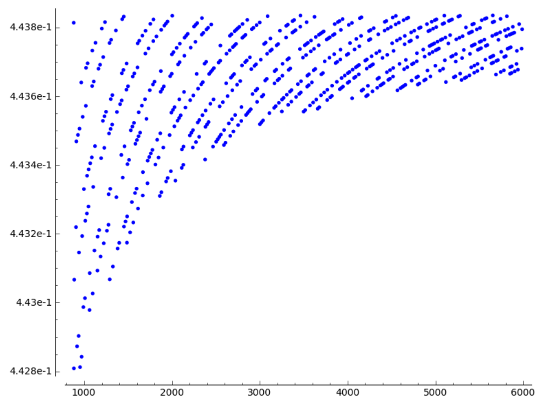
Upper bounds in the case $n = 1$. When we plot $a(N, 1)$ for $N < 300$, $N < 10000$ and $N < 30000$, we get



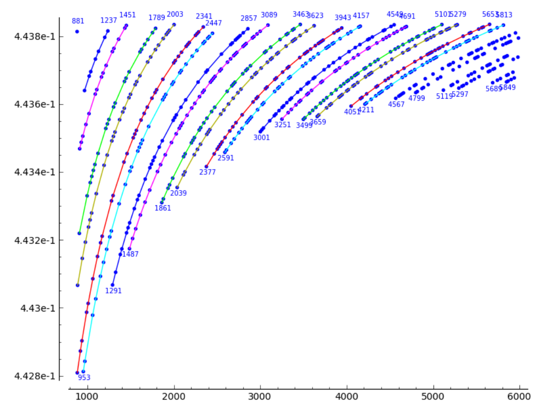
Note that we see a cloud of points but there are some outliers in the upper part. In the first plot, $N = 2, 5, 13, 89, 233$. One could think that this is an effect of the small values, but in the second plot we recover a “last” outlier, at $N = 1597$. The third plot shows that it was not the last, we have also $N = 28657$.

The explanation follows from our analysis of the upper bounds (3.36). This outliers are Fibonacci primes. Although it is not known if there are infinitely many of them, the exponential growth of the Fibonacci numbers explain why they are isolated.

A conjectural property for $n = 2$. When we plot $\big(N, a(N,2)\big)$ a clear pattern appears. The whole plot seems to be made of a kind of arcs of hyperbola

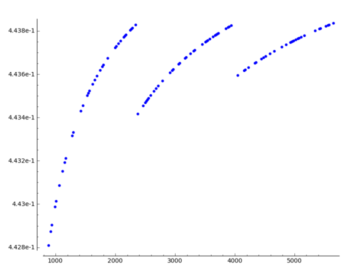


$880 \leq \mathrm{N} < 6000$

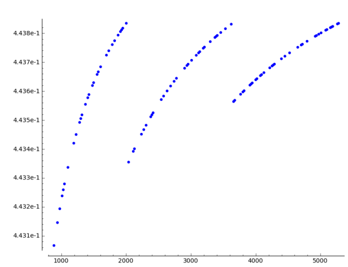


$880 \leq \mathrm{N} < 6000$

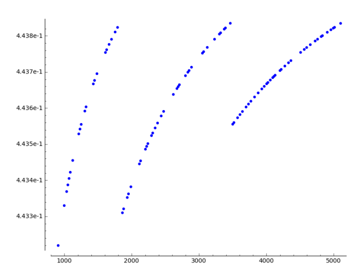
The numbers coprime to 18 are in the classes $r_1 = 7$, $r_2 = 5$, $r_3 = 1$, $r_4 = 17$, $r_5 = 13$, $r_6 = 11$. Then for any $N > 3$ prime $N \equiv r_j \pmod{18}$ for some r_j. Surprisingly, the pattern becomes very simple when subdividing N according to congruence classes modulo 18. For instance:



$880 \leq N < 5654$, $N \equiv 1$ (18)



$880 \leq N < 5280$, $N \equiv 5$ (18)



$880 \leq N < 5102$, $N \equiv 7$ (18)

The experimental data suggest *the existence of a sequence of primes*

$$N_1 < N_2 < N_3 < N_4 \dots \qquad \textit{with} \quad N_{6j+i} \equiv r_i \pmod{18} \quad \textit{for } i = 1, \dots, 6,$$

such that $a(N,2)$ is increasing for the primes $N \in (N_j, N_{j+6}]$ with $N \equiv N_j \pmod{18}$.

We have verified the conjecture for $N < 30000$. The sequence N_j $1 \le j \le 72$ is listed here:

j	N_j	j	N_j	j	N_j	j	N_j	j	N_j	j	N_j	j	N_j	j	N_j
1	151	10	2447	19	5101	28	7451	37	9907	46	12401	55	14947	64	17333
2	347	11	2857	20	5279	29	7789	38	10211	47	12757	56	15107	65	17599
3	631	12	3089	21	5653	30	7949	39	10567	48	12953	57	15427	66	17669
4	881	13	3463	22	5813	31	8377	40	10709	49	13309	58	15641	67	18223
5	1237	14	3623	23	6151	32	8573	41	11119	50	13487	59	16033	68	18401
6	1451	15	3943	24	6329	33	8929	42	11279	51	13807	60	16229	69	18793
7	1789	16	4157	25	6703	34	8999	43	11617	52	13967	61	16603	70	18917
8	2003	17	4549	26	6917	35	9463	44	11831	53	14341	62	16763	71	19309
9	2341	18	4691	27	7237	36	9623	45	12097	54	14591	63	17137	72	19469

If in the initial plot, we mark the arcs and we draw the different congruence classes with different colors, we can see the meaning of this conjecture. The upper extreme of the arcs in the right plot above correspond to N_j. Let M_j be the smallest prime number $M_j > N_j$ in the same congruence class, $18 \mid M_j - N_j$. These numbers are indicated in the lower extreme of the arcs.

We have also checked experimentally the following facts in the same range:

- In the computation of $a(N_j)$, the minimum is reached at the last but one convergent in the continued fraction. This implies $kp \equiv \pm 1 \pmod{N_j}$ with the original notation of the problem. The sign depends on the parity of the length of the continued fraction of k/N.
- The same applies in the computation of $a(M_j)$.
- It seems that $a(N_j)/a(M_j) = 1 + O\big(N_j^{-1}\big)$. Numerically, for $N_j < 30000$, it holds $\big|a(N_j)/a(M_j) - 1\big| < 2.23/N_j$.
- It seems that N_j and M_j grow more or less linearly on j.

Bibliography

[1] O. Aharony, S. S. Gubser, J. Maldacena, H. Ooguri, and Y. Oz. Large N field theories, string theory and gravity. *Phys. Rep.*, 323(3-4):183–386, 2000.

[2] I. J. R. Aitchison and A. J. G. Hey. *Gauge Theories in Particle Physics, Volume II. A Practical Introduction. Non-Abelian Gauge Theories. QCD and the Electoweak Theory (Graduate Student Series in Physics).* Taylor & Francis, 2003.

[3] I.J.R. Aitchison and A.J.G. Hey. *Gauge Theories in Particle Physics: Volume I: From Relativistic Quantum Mechanics to QED, Third Edition.* Graduate Student Series in Physics. CRC Press, 2002.

[4] G. Bhanot, U. M. Heller, and H. Neuberger. The quenched Eguchi-Kawai model. *Physics Letters B*, 113(1):47–50, 1982.

[5] M. Born and E. Wolf. *Principles of optics: Electromagnetic theory of propagation, interference and diffraction of light.* With contributions by A. B. Bhatia, P. C. Clemmow, D. Gabor, A. R. Stokes, A. M. Taylor, P. A. Wayman and W. L. Wilcock. Third edition. Pergamon Press, Oxford-New York-Paris, 1965.

[6] J. Bourgain and A. Kontorovich. On Zaremba's conjecture. *Ann. of Math. (2)*, 180:137–196, 2014.

[7] E. Brézin and S.R. Wadia, editors. *The large N expansion in quantum field theory and statistical physics: From spin systems to 2-dimensional gravity.* World Scientific Publishing Co. Pte. Ltd., 1993.

[8] N. Cabibbo and E. Marinari. A new method for updating $SU(N)$ matrices in computer simulations of gauge theories. *Phys. Lett. B*, 119(4-6):387–390, 1982.

[9] F. Chamizo and A. González-Arroyo. Tachyonic instabilities in 2+1 dimensional Yang-Mills and its connection to number theory results and conjectures. (In preparation).

[10] J. Cilleruelo and A. Córdoba. *La teoría de los números*. Biblioteca Mondadori. Mondadori España, Madrid, 1992.

[11] S. Coleman. $1/N$. In *Pointlike structures inside and outside hadrons*, Subnuclear series. Plenum Press, 1982.

[12] R. Courant and D. Hilbert. *Methods of mathematical physics. Vol. I*. Interscience Publishers, Inc., New York, N.Y., 1953.

[13] M. Creutz. Group theory and lattice gauge fields. In *Lectures presented at the Symposium on Lattice Gauge Theory Using Parallel Processors, Beijing, Chin, 21 May - 2 June*. 1987.

[14] S. R. Das. Some aspects of large-N theories. *Rev. Mod. Phys.*, 59(1), 1987.

[15] Ph. de Forcrand and O. Jahn. Monte Carlo overrelaxation for $SU(N)$ gauge theories. In *QCD and numerical analysis III*, volume 47 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, Berlin, 2005.

[16] P. A. M. Dirac. The quantum theory of the electron. *Proc. Royal Soc. London A*, 117(778):610–624, 1928.

[17] P. A. M. Dirac. The Lagrangian in Quantum Mechanics. *Physikalische Zeitschrift der Sowjetunion*, 3(1):610–624, 1933.

[18] T. Eguchi and H. Kawai. Reduction of dynamical degrees of freedom in the large-N gauge theory. *Phys. Rev. Lett.*, 48:1063–1066, Apr 1982.

[19] E. Elizalde. *Ten physical applications of spectral zeta functions*, volume 855 of *Lecture Notes in Physics*. Springer, Heidelberg, second edition, 2012.

[20] G. Feinberg. Possibility of faster-than-light particles. *Phys. Rev.*, 159:1089–1105, 1967.

[21] R. P. Feynman. An operator calculus having applications in quantum electrodynamics. *Physical Rev. (2)*, 84:108–128, 1951.

[22] R. P. Feynman. *QED: The Strange Theory of Light and Matter*. Princeton University Press, 1985.

[23] R. P. Feynman and A. R. Hibbs. *Quantum mechanics and path integrals*. McGraw-Hill, New York, 1965.

[24] G. B. Folland. *Quantum field theory*, volume 149 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2008. A tourist guide for mathematicians.

[25] T. Frankel. *The geometry of physics*. Cambridge University Press, Cambridge, third edition, 2012. An introduction.

[26] A. Galindo and P. Pascual. *Quantum mechanics. I*. Texts and Monographs in Physics. Springer-Verlag, Berlin, 1990. Translated from the Spanish by J. D. García and L. Alvarez-Gaumé.

[27] D. V. Gal′tsov, Yu. V. Grats, and V. Ch. Zhukovskiĭ. *Campos clásicos*. Editorial URSS, 2005.

[28] M. García Pérez, A. González-Arroyo, L. Keegan, M. Okawa, and A. Ramos. A comparison of updating algorithms for large N reduced models. *J. High Energy Phys.*, (6):193, front matter+20, 2015.

[29] M. García Pérez, A. González-Arroyo, and M. Okawa. Spatial volume dependence for $2+1$ dimensional $SU(N)$ Yang-Mills theory. *JHEP*, 2013(9), 2013.

[30] M. García Pérez, A. González-Arroyo, and M. Okawa. Volume independence for Yang-Mills fields on the twisted torus. *Int. J. Mod. Phys. A*, 29(1445001), 2014.

[31] M. Göckeler and T. Schücker. *Differential geometry, gauge theories, and gravity*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, Cambridge, 1987.

[32] A. González-Arroyo. Yang-Mills fields on the 4-dimensional torus (classical theory). In *Advanced School on Non-perturbative Quantum Field Physics (Peñiscola, 1997)*, pages 57–91. World Sci. Publ., River Edge, NJ, 1998.

[33] A. González-Arroyo and M. Okawa. The twisted-Eguchi-Kawai model: A reduced model for large-N lattice gauge theory. *Phys. Rev. D*, 27(10):2397, 1983.

[34] A. González-Arroyo and M. Okawa. A twisted model for large N lattice gauge theory. *Physics Letters B*, 120:174–178, jan 1983.

[35] A. González-Arroyo and M. Okawa. Large N reduction with the twisted Eguchi-Kawai model. *J. High Energy Phys.*, (7):043, 16, 2010.

[36] A. González-Arroyo and M. Okawa. The string tension for large N gauge theory from smeared Wilson loops. *PoS*, LATTICE2012:221, 2012.

[37] A. González-Arroyo and M. Okawa. Confinement in large N gauge theories. arXiv:1303.4921, 2013.

[38] J. F. González H. Physics and number theory: a romance with many polylogarithms and special functions [Spanish]. *Gac. R. Soc. Mat. Esp.*, 18(3):569–588, 2015.

[39] I. S. Gradshteyn and I. M. Ryzhik. *Table of integrals, series, and products*. Elsevier/Academic Press, Amsterdam, seventh edition, 2007. Translated from the Russian, Translation edited and with a preface by A. Jeffrey and D. Zwillinger.

[40] M. N. Huxley. *Area, lattice points, and exponential sums*, volume 13 of *London Mathematical Society Monographs. New Series*. The Clarendon Press, Oxford University Press, New York, 1996. Oxford Science Publications.

[41] E. Ibañez Bribán. A possible large N phase transition within the weak coupling regime of the twisted Eguchi-Kawai model. Master Thesis. Universidad Autónoma de Madrid – Instituto de Física Teórica UAM-CSIC, 2015.

[42] J. D. Jackson and L. B. Okun. Historical roots of gauge invariance. *Rev. Modern Phys.*, 73(3):663–680, 2001.

[43] A. Kontorovich. From Apollonius to Zaremba: local-global phenomena in thin orbits. *Bull. Am. Math. Soc., New Ser.*, 50(2):187–228, 2013.

[44] T. Lancaster and S. J. Blundell. *Quantum Field Theory for the Gifted Amateur*. Oxford University Press, New York, 2014.

[45] D. R. Lebedev and M. I. Polikarpov. Extrema of the twisted Eguchi-Kawai action and the finite Heisenberg group. *Nuclear Phys. B*, 269(2):285–294, 1986.

[46] T. Levi-Civita. *The absolute differential calculus*. Dover Phoenix Editions. Dover Publications, Inc., Mineola, NY, 2005. Reprint of the 1926 translation.

[47] B. Lucini and M. Panero. SU(N) gauge theories at large N. *Phys. Rep.*, 526(2):93–163, 2013.

[48] Y. Makeenko. A brief introduction to Wilson loops and large n. *Phys. Atom. Nucl.*, 73:878–894, 2010.

[49] S. J. Miller and R. Takloo-Bighash. *An invitation to modern number theory*. Princeton University Press, Princeton, NJ, 2006. With a foreword by Peter Sarnak.

[50] G. Münster and M. Walzl. Lattice gauge theory. A short primer. arXiv:hep-lat/0012005, 2000.

[51] S. P. Novikov and I. A. Taimanov. *Modern geometric structures and fields*, volume 71 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2006.

[52] M. Okawa. Monte Carlo study of the quenched Eguchi-Kawai model. *Phys. Rev. Lett.*, 49:705–707, Sep 1982.

[53] L. O'Raifeartaigh and N. Straumann. Gauge theory: historical origins and some modern developments. *Rev. Modern Phys.*, 72(1):1–23, 2000.

[54] M. E. Peskin and D. V. Schroeder. *An introduction to quantum field theory*. Addison-Wesley Publishing Company, Advanced Book Program, Reading, MA, 1995. Edited and with a foreword by David Pines.

[55] H. D. Politzer. Reliable perturbative results for strong interactions? *Phys. Rev. Lett.*, 30:1346–1349, 1973.

[56] H. Rademacher and O. Toeplitz. *The enjoyment of math*. Princeton Science Library. Princeton University Press, Princeton, NJ, 1994. Translated from the second (1933) German edition and with additional chapters by H. Zuckerman.

[57] H. J. Rothe. *Lattice gauge theories*, volume 74 of *World Scientific Lecture Notes in Physics*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, third edition, 2005. An introduction.

[58] J. Schwinger, editor. *Selected papers on quantum electrodynamics*. Dover Publications, Inc., New York, 1958.

[59] C. L. Siegel. Contributions to the theory of the Dirichlet L-series and the Epstein zeta-functions. *Ann. of Math. (2)*, 44:143–172, 1943.

[60] C. L. Siegel. *Lectures on advanced analytic number theory*. Notes by S. Raghavan. Tata Institute of Fundamental Research Lectures on Mathematics, No. 23. Tata Institute of Fundamental Research, Bombay, 1965.

[61] J. Smit. *Introduction to quantum fields on a lattice*, volume 15 of *Cambridge Lecture Notes in Physics*. Cambridge University Press, Cambridge, 2002.

[62] M. Spivak. *A comprehensive introduction to differential geometry. Vol. II*. Publish or Perish, Inc., Wilmington, Del., second edition, 1979.

[63] H. E. Stanley. Spherical model as the limit of infinite spin dimensionality. *Physical Review*, 176:718–721, dec 1968.

[64] G. 't Hooft. A planar diagram theory for strong interactions. *Nuclear Phys. B*, 72:461–473, 1974.

[65] G. 't Hooft. A two-dimensional model for mesons. *Nucl. Phys.*, B75:461–470, 1974.

[66] G. 't Hooft. A property of electric and magnetic flux in non-Abelian gauge theories. *Nuclear Phys. B*, 153(1-2):141–160, 1979.

[67] G. 't Hooft. Aspects of quark confinement. *Phys. Scripta*, 24(5):841–846, 1981.

[68] G. 't Hooft. Some twisted self-dual solutions for the Yang-Mills equations on a hypertorus. *Comm. Math. Phys.*, 81(2):267–275, 1981.

[69] G. 't Hooft. *Under the spell of the gauge principle*, volume 19 of *Advanced Series in Mathematical Physics*. World Scientific Publishing Co., Inc., River Edge, NJ, 1994.

[70] G. 't Hooft. Large N. arXiv:hep-th/0204069, 2002.

[71] M. Teper and H. Vairinhos. Symmetry breaking in twisted Eguchi-Kawai models. *Phys. Lett. B*, 652(5-6):359–369, 2007.

[72] A. Terras. *Fourier analysis on finite groups and applications*, volume 43 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1999.

[73] P. van Baal and B. van Geemen. A simple construction of twist-eating solutions. *J. Math. Phys.*, 27(2):455–457, 1986.

[74] S. Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.

[75] K.G. Wilson. Confinement of quarks. *Phys. Rev. D*, 10(8):2445–2459, 1974.

[76] E. Witten. Baryons in the $1/N$ expansion. *Nuclear Phys. B*, 160:57–115, 1979.

[77] C. N. Yang. Conceptual origin of Maxwell equations and of gauge theory. ASIAA/CCMS/IAMS/LeCosPA/NTU-Phys Joint Colloquia. `https://www.youtube.com/watch?v=SG343kojbnU`, 2015.

[78] C. N. Yang and R. L. Mills. Conservation of isotopic spin and isotopic gauge invariance. *Physical Rev. (2)*, 96:191–195, 1954.

[79] A. Zee. *Quantum field theory in a nutshell*. Princeton University Press, Princeton, NJ, 2003.

[80] D. Z. Zhang. C. N. Yang and contemporary mathematics. *Math. Intelligencer*, 15(4):13–21, 1993.