Notes on the analytic S-matrix (under construction)

Alexander Zhiboedov

CERN, Theoretical Physics Department, 1211 Geneva 23, Switzerland

Abstract

We review the basic notions of S-matrix theory. We discuss the status of the nonperturbative S-matrix bootstrap program. We focus on topics and methods where we believe a further progress can be readily made.

Contents

1 Introduction

"One is never sure to have completely exploited the axioms of QFT."
A. Martin

"We shall be content with plausibility arguments."
R. Eden, P. Landshoff, D. Olve, J. Polkinghorne, "The analytic S-matrix."

In these notes, we will discuss the basic properties of *nonperturbative* relativistic scattering amplitudes. Understanding these is the subject of S-matrix theory.

S-matrix theory grew to prominence at the end of the fifties (see, e.g. [1] for a historical review). It has been realized that combining special relativity and quantum mechanics possesses certain rigidity (bootstrapiness), which gave hope to the idea of solving relativistic quantum field theories based on self-consistency plus a little more. "A little more" could be an educated guess, experimental data or the result of trial-and-error.

This first attempt to bootstrap the S-matrix was not successful because it was not possible to answer two basic questions:

- What is a complete set of fundamental principles upon which S-matrix theory should be based?
- What are the reliable computational schemes that allow *systematically* going from basic principles to interesting physical results?

After the end of the sixties, the nonperturbative S-matrix theory was largely dormant thanks to the discovery of string theory and QCD (the former being the outcome of S-matrix theory itself). The development of these theories left behind abandoned, beautiful S-matrix constructions with many explorations unfinished and many directions untouched. Another reason for moving on was that the subject has proven to be very difficult.

Since then we have experienced many developments in a few related areas. Let us list the important ones:

1. A revolution in our ability to compute perturbative amplitudes due to a plethora of new on-shell methods and smart mathematical tricks.[1] These methods, very much in spirit of the original S-matrix program, are perturbative in nature and so far have not shed new light on the nonperturbative amplitudes.
2. Discovery of the AdS/CFT correspondence. From the S-matrix bootstrap point of view, AdS/CFT answers the first question above in AdS: the fundamental principles upon which the scattering in AdS is built are CFT axioms.
3. Development of the conformal bootstrap. The conformal bootstrap answers the second questions above and offers a set of computational tools to get interesting results starting from the CFT axioms.

[1] As S. Weinberg quipped "quantum field theory is S-matrix theory, made practical."

In other words, AdS/CFT + Conformal Bootstrap fully realize the original dream of S-matrix theory in AdS. AdS/CFT was a leap of imagination, and in our opinion this is also what is needed to make progress in the S-matrix bootstrap.

With these insights, we face the same question as half of a century before: what are the general principles for scattering relativistic particles in flat space? How constraining are they? There is a hope that we can make further progress. We can use experience from the conformal bootstrap, AdS/CFT, and new perturbative methods. We can exploit much more powerful computational resources (old ideas shine on new computers). Last but not least, we can have a much more open spirit of investigation (we do not have to solve the actual QCD! There are other interesting theories out there).

In fact, the main message of these lectures is: we have a reasonable, good guess for what the rules of the game are, and there are some methods available to utilize them. In other words, there is a lot of work to be done!

The subject has been already experiencing a new life recently. New efforts were made to bootstrap scattering amplitudes of strongly coupled massive QFTs, see e.g. [29, 32], as well as gravitational theories (see e.g. [33] and references therein). With these comments in mind, we try to review below old results and more recent developments based on which we can make further progress. Hopefully, this little note can be helpful for students and researchers who either would like to get a glimpse of this beautiful subject or, even better, start working on it.

1.1 Literature

The literature on the subject is vast. Very often thinking that we have found something new, we later discovered another book or article that proved otherwise. We expect that this natural process of rediscovering things will end in the next few years.

Let us list below a few books and review articles that we found particularly useful.

- R. Eden, P. Landshoff, D. Olve, J. Polkinghorne, "The analytic S-matrix."
  This book is classics but not an easy read. Chapters 1 and 4 serve as an excellent introduction to the basics. The book's core is chapter 2, which one could read in the stream-of-consciousness mode because, in some sense, it ends nowhere (hopefully, after these lectures, things will become clearer). With some effort, it does contain a lot of exciting material and covers a review of an unfinished attempt to prove the Mandelstam representation. It also serves as a good review of what has been achieved on the subject by the Cambridge group. For example, to the best of our knowledge, the analogous results achieved by the Soviet group (Landau, Gribov, Kolkunov, Okun, Patashinski, Petrina, Rudik, Sudakov,...) are not adequately summarized anywhere. The original articles are not even to be found via Inspire.

- G. Chew, "The analytic S-matrix."
  A good source to read about philosophy (if not ideology) behind the old S-matrix bootstrap. It contains a detailed review of nuclear democracy[2] and the strip approximation, "whose mention will bring tears to the eyes of those of us who are old enough to remember it" (S. Weinberg).

[2]This is a postulate that analyticity in spin includes $J = 0$ partial waves. When applied to the OPE data, is "operator democracy" realized in the 3d Ising (general CFTs)?

- R. Eden, "High Energy Collisions of Elementary Particles."
  In addition to the standard material, chapter 2 reviews some basic experimental data obtained in the (old) collider experiments which is sometimes used as a guide for making various theoretical assumptions.

- P. Collins, "An introduction to Regge theory and High-Energy Physics."

- P. Collins, E. Squires, "Regge Poles in Particle Physics."
  These two books contain an extensive review of various results related to the physics of complex angular momentum.

  Let us now mention a few review articles and lecture courses:

- G. Sommer, "Present State of Rigorous Analytic Properties of Scattering Amplitudes."
  The most systematic and user-friendly review of the derivation of classic results of analyticity in QFT known to us.

- A. Martin, F. Cheung "Analyticity Properties and Bound of the Scattering Amplitudes."
  Another source with a particular emphasis on the unitarity extension of the analyticity domain.

- R. Eden, "Theorems on High Energy Collisions of Elementary Particles."
  An incredibly dense four page collection of various bounds with relevant references and assumptions used.

A few other useful books:

- D. Iagolnitzer, "Scattering in Quantum Field Theory."

- V. Gribov, "Theory of complex angular momentum."

- V. Gribov, Y. Dokshitzer, J. Nyiri, "Strong interactions of hadrons at high energies."

- S. Donnachie, G. Dosch, P. Landshoff, O. Nacthmann, "Pomeron Physics and QCD."

Last but not the least recently there has been a series of lectures on the nonperturbative aspects of the S-matrix bootstrap by Balt van Rees, Pedro Vieira, Simon Caron-huot, and Andrea Guerrieri (all available online), which we hope these lectures can complement. A lot of the material in the notes is based on a joint work with Amit Sever and Miguel Correia [35], who I also would like to thank for many discussions on the subject.

1.2 Plan

Our (tentative!) plan is the following:

1. Introduction: general overview, the signal model.

2. Relativistic kinematics, *unitarity*, partial wave expansion.

3. *Analyticity* and *crossing*: axiomatic analyticity, Landau equations, maximal analyticity.

4. Dispersion relations, bound on chaos, superconvergence, null constraints.

5. Nonperturbative analytic bootstrap: the Froissart-Gribov formula, Aks theorem, Dragt bootstrap.

6. Nonperturbative numerical bootstrap: scattering amplitudes as unitarity map fixed points, primal problem, dual problem.

2 Signal Model: Causality, Analyticity, Unitarity

The basic question that is asked in S-matrix theory is the following: what is a transition amplitude between a given state of free particles in the far past and in the far future? All such transition amplitudes are encoded in a unitary operator, the S-matrix, that contains all information about the particle interactions.

In particular, the spacetime in this picture is emergent. One natural question is then: how is *causality* encoded in the properties of scattering amplitudes? This is a nontrivial question to which a complete answer is not known. Nevertheless this question naturally leads to one of the pillars of S-matrix theory, namely *analyticity*.

Another important property of the S-matrix is *unitarity* which is an expression of quantum-mechanical nature of the theory.

To get some intuition behind the various notions of S-matrix theory, it is instructive to consider a toy model which nevertheless retain some of the basic properties of relativistic amplitudes. The toy model we are about to discuss is usually called the signal model, see appendix D of [34].[3]

We consider an initial signal which is a function of time[4] $f_{\rm in}(t)$ and an out-signal $f_{\rm out}(t)$. We postulate that the two are related as follows

$$f_{\rm out}(t) = \int_{-\infty}^{\infty} dt' S(t-t') f_{\rm in}(t') \ , \tag{1}$$

where $S(t-t')$ is the S-matrix whose properties we would like to understand. In the context of relativistic scattering we can think about scattering of a given particle off a background of other particles (their number can change in the process of interaction).

Switching to the Fourier space we get

$$S(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega t} S(\omega) \tag{2}$$

and the scattering relation takes the form[5]

$$f_{\rm out}(\omega) = S(\omega) f_{\rm in}(\omega) \ . \tag{3}$$

Causality is the statement that if $f_{\rm in}(t') = 0$ for $t' < 0$, then $f_{\rm out}(t) = 0$ for $t < 0$ or, more generally, that the outgoing signal $f_{\rm out}(t)$ depends on $f_{\rm in}(t')$ for $t' \leq t$ only. Through (1) this implies that $S(t) = 0$ for $t < 0$. Let us see what it implies for its Fourier transform $S(\omega)$

$$S(\omega) = \int_0^{\infty} dt e^{i\omega t} S(t). \tag{4}$$

[3] See also Simon Caron-Huot's notes from the Bootstrap school 2020.

[4] In the relativistic case it will be the null (or retarded) time.

[5] In the relativistic case we will write it as $\delta(\omega-\omega')T(\omega)$ to make energy conservation (or time translation) manifest.

It is immediately clear from (4) that we can continue $S(\omega)$ to the upper half-plane. Indeed $\text{Im}[\omega] > 0$ only improves the convergence of the integral.

It is therefore natural to model physical S-matrices by considering function $S(\omega)$ which is *analytic* in the upper half-plane. We can then define the physical S-matrix to be a limit of the analytic function from the upper half-plane

$$S(\omega) = \lim_{\epsilon \to 0} S(\omega + i\epsilon) \ . \tag{5}$$

This is known as $i\epsilon$ prescription. Below we will see how a related statement holds for the actual relativistic scattering amplitudes.

It is crucial, that analyticity in the upper half-plane while being *necessary* is not *sufficient* for causality.

Exercise: Show that $S(\omega) = e^{i\omega^3}$ while being analytic in the upper half-plane violates causality. Time delay in gravity originates from $S(\omega) = e^{i\alpha\omega}$. Check that causality implies that $\alpha \geq 0$.

To understand the sufficient condition we consider (2) and try to use analyticity in the upper half-plane to establish that $S(t) = 0$ for $t < 0$. The standard argument for this involves deformation of the contour into the upper half-plane to argue that the integral vanishes due to the factor $e^{t\text{Im}[\omega]}$ factor. This however assumes that $|S(\omega)|$ grows slower than exponential in the upper half-plane. The conclusion is that $S(\omega)$ describes causal transfer if and only if it is analytic in the upper half-plane, $\text{Im}\ \omega > 0$, and *grows slower than exponential* when $\text{Im}\ \omega \to \infty$.

In many physically interesting cases a stronger condition is known to hold (these include unitary relativistic QFTs and gravitational theories). The physical S-matrices are in fact polynomially bounded in the upper half-plane

$$|S(\omega)| < |\omega|^N, \quad |\omega| \gg 1 \ . \tag{6}$$

The precise value of N will depend on the details of the theory (massive, massless, gravitational or not, dimensionality of spacetime d). In all cases however the extra input that leads to a stronger bound (6) comes from *unitarity.*

2.1 Unitarity

Unitarity expresses the fact that scattering describes time evolution in a quantum theory. Therefore we can introduce transition probabilities and these should sum up to 1. In the context of the signal model, where by assumption the initial and final states describe the same particle, the probability of the transition through the box should be ≤ 1. In other words, we do not want the black box to enhance the signal.

By unitarity in the signal model we mean the following: the L_2 norm of the out-signal should be smaller than that of the in-signal

$$||f_{\text{out}}||^2 = \int dt |f_{\text{out}}(t)|^2 \leq \int dt |f_{\text{in}}(t)|^2 = ||f_{\text{in}}||^2. \tag{7}$$

In other words we would like to impose that the scattering process does not enhance $||...||^2$.

We will now prove that unitarity implies that

$$|S(\omega)| \leq 1, \quad \text{Im}[\omega] \geq 0. \tag{8}$$

At large $|\omega|$ this is the same as (6) with $N = 0$. In the relativistic setup we will encounter cases with $N > 0$.

With some foresight, we pick a particular $f_{\rm in}(t)$ of the form

$$f_{\rm in}(t) = e^{-\gamma t} e^{-i\omega_0 t}\theta(t)\sqrt{2\gamma}\ , \quad f_{\rm in}(\omega) = \frac{\sqrt{2\gamma}\ i}{\omega - \omega_0 + i\gamma}. \tag{9}$$

with $\gamma > 0$ and ω_0 real. Note that $||f_{\rm in}||^2 = \int dt |f_{\rm in}(t)|^2 = 1$.

For $\text{Im}(\omega) > 0$ we can now write

$$\begin{aligned} |f_{\rm out}(\omega)|^2 &\leq \left| \int_0^\infty dt e^{i\omega t} f_{\rm out}(t) \right| \leq \int_0^\infty dt |e^{i\omega t}|^2 \int_0^\infty dt |f_{\rm out}(t)|^2 = \frac{1}{2\text{Im}(\omega)} ||f_{\rm out}||^2 \\ |f_{\rm out}(\omega)|^2 &\leq \frac{1}{2\text{Im}(\omega)} \\ |S(\omega)|^2 &= \frac{|f_{\rm out}(\omega)|^2}{|f_{\rm in}(\omega)|^2} \leq \frac{1}{2\text{Im}(\omega)} \frac{1}{|f_{\rm in}(\omega)|^2}\ . \end{aligned} \tag{10}$$

Here we used the Cauchy-Schwartz inequality. Note that $|e^{i\omega t}|^2 = e^{-2\text{Im}(\omega)t}$. We also used that $||f_{\rm out}||^2 \leq ||f_{\rm in}||^2 = 1$. We can now set $\omega = \omega_0 + i\gamma$ and find that $f_{\rm in}(\omega_0 + i\gamma) = 1/\sqrt{2\gamma}$ for the specific function (9). Inserting this into the equation above we then find that $|S(\omega_0 + i\gamma)| \leq 1$, which is what we wanted to prove, since ω_0 and γ are arbitrary.

In conclusion, we find that $S(\omega)$ should be analytic and bounded $|S(\omega)| \leq 1$ in the upper half plane. These are necessary and sufficient conditions for causality and unitarity in the signal model that we considered.

Exercise: Consider $S(\omega) = 1 + ic\omega^p + O(c^2)$, where $c \ll 1$. What are the bounds on c and p coming from $|S(\omega)| \leq 1$ for $\text{Im}\omega \geq 0$. The condition that $p \leq 1$ is also known as the *bound on chaos*. It is saturated in the classical theory of gravity and manifests itself as the Shapiro time delay in physical experiments.

2.2 Relation to the Relativistic Scattering: coherent states

The argument above is very suggestive. It is desirable to translate it to a rigorous bound on the $2 \to 2$ scattering amplitude of relativistic particles. An immediate problem is that for particles we have $\omega > 0$ only. Let us reprint the path taken in [34].

We consider light cone coordinates u and v. We consider a perturbation that is translation invariant in v but is localized in the u coordinate. We call this "the shockwave background" (it can be a relativistic particle with small p_v). We expand the fields in the v coordinate at some $u_{\rm in}$ and then we expand them again at some $u_{\rm out}$ after the shock. To make contact with the above discussion we take $t = v$ and $p_v = -\omega$. We can expand the field as

$$\phi(t) = \int_0^\infty \frac{d\omega}{\sqrt{\omega}} (a_\omega e^{-i\omega t} + a_\omega^\dagger e^{i\omega t})\ . \tag{11}$$

We should think of this field as an asymptotic free field. We next define

$$S(\omega) = -\int_{-\infty}^\infty dt e^{i\omega t} [\phi_{\rm out}(t), i\partial_t \phi_{\rm in}(0)] = -\int_0^\infty dt e^{i\omega t} [\phi_{\rm out}(t), i\partial_t \phi_{\rm in}(0)]\ , \tag{12}$$

where $\phi_{\rm in}(t)$ is the field operator before the interaction, say $u \le u_{in}$, and $\phi_{\rm out}(t)$ is the field operator after the interaction ($u \ge u_{out}$). With this definition we have

$$S(-\omega) = S^*(\omega). \tag{13}$$

If we consider a signal that is made out of physical particles, one might correctly worry that the fact that $\omega > 0$ will preclude us from localizing the signal in time. In order to avoid this issue we can consider a coherent state of the form

$$|\Psi\rangle = e^{i\int dt f_{\rm in}(t)\phi_{\rm in}(t)}|0\rangle \ , \tag{14}$$

with a real function $f_{\rm in}$. This is a state that could be produced by adding a hermitian term to the Hamiltonian at some early time $u_{\rm in}$. On this state we have the expectation values

$$\langle\Psi|\partial_t\phi_{\rm in}(t)|\Psi\rangle = f_{\rm in}(t) \ , \qquad \langle\Psi|\partial_t\phi_{\rm out}(t)|\Psi\rangle = f_{\rm out}(t) \ , \tag{15}$$

where the functions are related as in the signal model. Here we assumed a linear relation between the in- and out-signals. Furthermore, we can also consider the expectation values of the normal ordered product $T_{vv} = T_{tt} =: \partial_t\phi(t)\partial_t\phi(t) :$. When this is evaluated on the state above, and integrated over t we find that the answer is given by

$$-P_v^{\rm in,out} = \int dv T_{vv} = \int dt (f_{\rm in,out}(t))^2 = ||f_{\rm in,out}||^2 \ . \tag{16}$$

Thus, the condition that the total light-cone momentum P_v should not increase implies the norm condition $||f_{\rm out}||^2 \le ||f_{\rm in}||^2$.

More precisely, we can consider the signal $f_{\rm in}$ exciting a mode involving a graviton with a given polarization. The signal $f_{\rm out}$ is the same mode of the graviton. In addition, the initial graviton could go into other massive particles. Then the condition that the total P_v in the out-graviton mode should be no bigger than the initial P_v, which was all contained in the graviton mode, leads to the norm condition (or unitarity condition) for the signal model. In conclusion, the graviton-graviton matrix element $S_{gg}(\omega)$ obeys all the assumptions of the signal model. Therefore, it should be analytic and $|S_{gg}(\omega)| \le 1$ in the upper half plane.

Let us now try to relate the picture above to the $2 \to 2$ scattering amplitude. The idea is that close to positive physical s we have

$$S(\omega) = S(s \sim \omega, b)\left(1 + O(\frac{1}{sb^2})\right), \quad sb^2 \gg 1, \tag{17}$$

where the correction comes from the fact that the the shockwave background when created by other particles carries a non-zero P_v momentum. We get $\frac{\delta p_v}{p_v} \sim \frac{q^2}{p_u p_v} \sim \frac{1}{sb^2} \ll 1$ for the argument to make sense.

One thing is clear from this argument: it would wonderful to make it directly in terms of the $2 \to 2$ scattering amplitude and avoid talking about coherent states, etc. This is what we discuss next.

2.3 Relation to the Relativistic Scattering: $2 \to 2$ scattering amplitude

The function $S(\omega)$ above captures two important features of the relativistic scattering amplitude

- analyticity
- unitarity

It misses however an important property of *crossing*.

At this point let us simply note that if we want to associate ω from the previous section with energy of a particle, we have to restrict it to $\omega > 0$. For $\omega > 0$ there is indeed a rather direct correspondence between $S(\omega)$ and $S(s, b)$ - the $2 \to 2$ scattering amplitude at fixed impact parameter b, where $s \sim \omega\omega'$.

As we will see in more detail below, $S(s, b)$ has similar properties: it is analytic for Im $s > 0$ and it satisfies unitarity on the right cut[6]

$$|S(s, b)| \leq 1 + O\left(\frac{1}{s^{1/2}b}\right), \quad s \geq 4m^2. \tag{18}$$

However, something new happens for $s < 0$ (which is analogous to $\omega < 0$): it describes a different scattering process (scattering in the crossed channel). Moreover, on the left cut we do not have a bound

$$|S(s, b)| \leq \ ?, \quad s < 0. \tag{19}$$

This is what prevents from turning the signal model argument into a complete relativistic argument.

Problem: Complete the argument above. Imagine we have a weakly coupled gravitational theory. Then the signal model suggests the following picture for $S(s, b)$, see Fig. 3. What happens on the left cut?

In fact only recently it has been understood how to derive the constraints of [34] systematically. Essentially, one uses *crossing* to solve the problem above with the left cut. To do it one goes back from the impact parameter space to the momentum space and implements crossing there. After it is done, one performs the impact parameter transform. Presumably, these developments should imply something about the left cut of $S(s, b)$, however this is still to be understood.

2.4 Comments

Relativistic particles are not something that can be easily localized. Trying to do so creates new particles! So what do we mean by causality then in a relativistic theory? There are two approaches to it.

Microcausality

In QFT we do have a sharp notion of causality when applied to local operators

$$[\mathcal{O}(x_1), \mathcal{O}(x_2)] = 0, \quad (x_1 - x_2)^2 - \text{spacelike}. \tag{20}$$

Going from this to the analytic properties of the scattering amplitude requires some further gymnastics, mainly the LSZ reduction and analytic completion. These steps are reviewed in great detail in the article by Sommer [5], and we refer the reader there for details. A further improvement of the QFT analyticity domain can be made by combining the arguments above with unitarity [3] (exactly in the spirit of the improvement due tounitarity above).

[6]The correction is related to the fact $J \sim \sqrt{s}b \gg 1$.

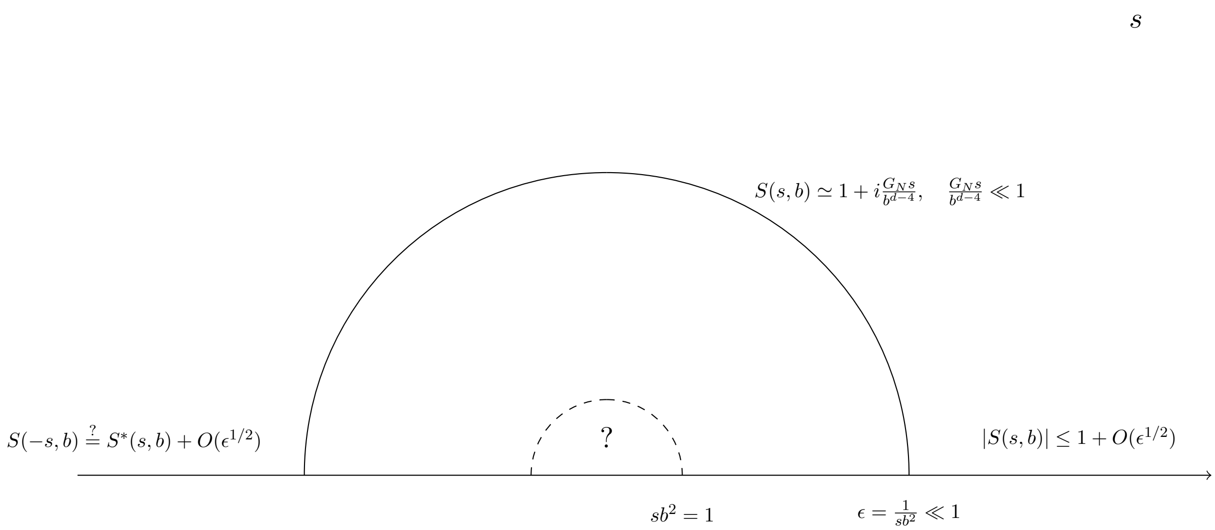

Figure 1: The structure of the s-plane for the phase shift $S(s,b)$. We consider weakly coupled gravitational theories such that on the ϵ-circle the phase shift is well approximated by the tree-level result $S(s,b) \simeq 1 + i\frac{G_N s}{b^{d-4}}$. We then can map the outside of the ϵ-disc and the upper half-plane to the unit disc and apply to it the maximum principle and the Schwarz-Pick lemma. The bound applies to terms that are parametrically larger than ϵ.

Macrocausality

A notion of causality that does not refer to local operators and is formulated in terms of particles directly is known as *macrocausality.* For a detailed review of this notion, see [18]. The basic idea is that information on macroscopic distances can only be transferred by on-shell particles and transfers in the disallowed regions are exponentially suppressed. Using macrocausality one can then derive certain analytic properties of the scattering amplitudes.

Problem: Formulate and work out the consequences of macrocausality in quantum gravity.

3 Relativistic Kinematics

In this section we review kinematics of relativistic scattering. In these lectures we will mostly discuss $2 \to 2$ scattering of scalar particles. In the next section we will discuss analytic properties of relativistic scattering amplitudes.

3.1 Particles

We consider a relativistic theory in d-dimensional Minkowski spacetime with spacetime coordinates $x^\mu = (t, \vec{x})$ and metric

$$ds^2 = -dt^2 + d\vec{x}^2 = \sum_{\mu,\nu=0}^{d-1} \eta_{\mu\nu} dx^\mu dx^\nu \ . \tag{21}$$

We assume that the theory is Poincare invariant and that at very early and late times the physical state is described by a system of free particles.

Particles are labeled by unitary, irreducible representations of the symmetry group, the mass m, and the spin $\mathbf{J}$ or, more precisely, the representation under the little group (for massive particles $SO(d-1)$; for massless particles $SO(d-2)$). For simplicity we restrict ourselves to theories of a single type of scalar particle (it transforms in the trivial representation of $SO(d-1)$). Scalar particles obey Bose statistics.

We will consider scattering of particles with definite momentum p^μ. The momentum of a particle satisfies the on-shell condition

$$-p^2 = (p^0)^2 - \vec{p}^2 = m^2, \quad p^0 > 0. \tag{22}$$

We assume that the theory has a Poincare invariant vacuum $|\Omega\rangle$. To describe the Hilbert space of the theory it is useful to introduce annihilation-creation operators. These obey canonical commutation relations

$$[a(\vec{p}), a^\dagger(\vec{q}] = (2\pi)^{d-1} 2p^0 \delta^{d-1}(\vec{p}-\vec{q}) \ . \tag{23}$$

The annihilation operator $a(\vec{p})$ acts on the vacuum as follows

$$a(\vec{p})|\Omega\rangle = 0. \tag{24}$$

The creation operator $a^\dagger(\vec{p})$ generates the one-particle state

$$|p\rangle \equiv a^\dagger(\vec{p})|\Omega\rangle \ . \tag{25}$$

The one-particle Hilbert space $\mathcal{H}_1$ is given by the space of square integrable functions $f(p)$ restricted to the mass-shell (22). The Lorentz-invariant scalar product in $\mathcal{H}_1$ is defined as:

$$\langle f_1|f_2\rangle = \int d\mu_1(p) f_1^*(p) f_2(p) \ ,$$
$$d\mu_1(p) \equiv \frac{1}{(2\pi)^{d-1}} \theta(p^0)\delta(p^2+m^2) d^d p, \tag{26}$$

where $|f\rangle = \frac{1}{(2\pi)^{d-1}} \int d^d p \ \theta(p^0)\delta(p^2+m^2) f(p) a^\dagger(\vec{p})|\Omega\rangle$. In a theory of a single scalar particle, the n-particle Hilbert space is given by a symmetrized product of one-particle Hilbert spaces $\mathcal{H}_n = (\mathcal{H}_1)^{\otimes_{\rm sym} n}$.

The full Hilbert space of the theory $\mathcal{H}$ is described by the Fock space of multi-particle states, which is a direct sum of n-particle states $\mathcal{H} = \oplus_{n=1}^\infty \mathcal{H}_n$ (with the condition that $||\phi|| < \infty$, where $||\phi|| = \sum_n ||\phi_n||$). We therefore have the following completeness relation

$$\hat{1} = |\Omega\rangle\langle\Omega| + \sum_{n=1}^\infty \int d\mu_n(p) |p_1, ..., p_n\rangle\langle p_1, ..., p_n| \ ,$$
$$d\mu_n(p) \equiv \frac{1}{n!} \prod_{i=1}^n \frac{d^d p_i}{(2\pi)^d} \theta(p_i^0)\delta(p_i^2+m^2). \tag{27}$$

3.2 The S-matrix

The S-matrix describes the evolution of a given state $|\psi_{\rm in}\rangle$ in the infinite past into a state in the infinite future $|\psi_{\rm out}\rangle$ (as a result of scattering). In other words, the S-matrix operator maps states to other states

$$\hat{S}: \quad |\psi_{\rm in}\rangle \to |\psi_{\rm out}\rangle \equiv \hat{S}|\psi_{\rm in}\rangle. \tag{28}$$

Due to its physical meaning it preserves norm of a state

$$\langle \psi_{\text{out}} | \psi_{\text{out}} \rangle = \langle \psi_{\text{in}} | \psi_{\text{in}} \rangle. \tag{29}$$

As a result, $\hat{S}$ is a unitary operator

$$\hat{S}\hat{S}^\dagger = \hat{S}^\dagger \hat{S} = \hat{1}. \tag{30}$$

We will be interested in probabilities of a given state $|\psi_1\rangle$ in the past to evolve into a given state $|\psi_2\rangle$ in the future. The transition amplitude that describes this process $\langle \psi_2 | \hat{S}\psi_1 \rangle$ is a complex number. Its modulus square $|\langle \psi_2 | \hat{S}\psi_1 \rangle|^2$ describes the transition probability.

For the m initial and n final particles we define the momentum space kernel $S_{m\to n}(q_1, ..., q_n; p_1, ..., p_m)$ so that for the corresponding states we have

$$\langle \psi_n | \hat{S}\psi_m \rangle = \int S_{m\to n}(q;p)\bar{\psi}_n(q_1, ..., q_n)\psi_m(p_1, ..., p_m) d\mu_m d\mu_n. \tag{31}$$

One can argue that $S_{m\to n}(q;p)$ are tempered distribution in the space of all on-shell momenta, see appendix I in [18].

Next we introduce the notion of the connected S-matrix using the idea of the spacetime cluster property which states that well-separated collections of particles do not interact with each other. This should hold in a gapped theory and it is an interesting question to what extent it holds in theories with massless particles. Mathematically, it means that

$$S_{m\to n} = S^c_{m\to n} + \sum_\pi \prod_k S^c_{m_k \to n_k} \ . \tag{32}$$

where the sum runs over nontrivial partitions of initial and final particles into subsets π_k. The clustering property (32) realizes the intuition that scattering between the states spatially displaced from each other factorizes as the relative displacement tends towards infinity.

Stability of particles imply that $S^c_{1\to k\geq 2} = 0$, as well as

$$S^c_{1\to 1} = 2(\vec{p}^2 + m^2)^{1/2}(2\pi)^{d-1}\delta^{d-1}(\vec{p} - \vec{q}) \ . \tag{33}$$

Finally, we can define scattering functions $T_{m,n}$ by removing the overall momentum-conservation δ-function from the connected S-matrix

$$S^c_{m\to n} = (2\pi)^d \delta^d \left(\sum_{i=1}^m q_i - \sum_{j=1}^n p_j \right) T_{m\to n}. \tag{34}$$

Unitarity of the S-matrix implies infinitely many nontrivial relations on $T_{m,n}$ also known as bubble diagrams.

In this note we are interested in the two-to-two scattering of identical particles of mass m. In this case we write

$$\begin{aligned} S_{2\to 2}(p_3, p_4|p_1, p_2) &= S^c_{1,1}(p_3, p_1)S^c_{1,1}(p_4, p_2) + S^c_{1,1}(p_4, p_1)S^c_{1,1}(p_4, p_2) + S^c_{2,2}(p_3, p_4|p_1, p_2), \\ S^c_{2,2}(p_3, p_4|p_1, p_2) &= i(2\pi)^d \delta^d \left(p_1 + p_2 - p_3 - p_4\right) T(s,t) \ , \end{aligned} \tag{35}$$

where we used Lorentz invariance to introduce Mandelstam invariants[7]

$$\begin{aligned} s &= -(p_1+p_2)^2 = 4(m^2+\mathbf{k}^2), \\ t &= -(p_1-p_3)^2 = 2\mathbf{k}^2(\cos\theta-1), \\ u &= -(p_1-p_4)^2 = 4m^2-s-t = 2\mathbf{k}^2(-\cos\theta-1) \ , \end{aligned} \tag{36}$$

where we evaluated them in the center-of-mass frame as well and introduced the scattering angle

$$\cos\theta = 1 + \frac{2t}{s-4m^2} \ . \tag{37}$$

Mandelstam invariants satisfy

$$s+t+u = 4m^2 \ . \tag{38}$$

From above we also have

$$[T(s,t)] = m^{4-d} \ , \tag{39}$$

for the amplitude dimensionality.

Based on the kinematics above we can talk about three channels:

$$\begin{aligned} s-\text{channel}: &\quad 1,2\to 3,4 \qquad s\geq 4m^2, \quad 4m^2-s\leq t\leq 0, \\ t-\text{channel}: &\quad 1,3\to 2,4 \qquad t\geq 4m^2, \quad 4m^2-t\leq s\leq 0, \\ u-\text{channel}: &\quad 1,4\to 2,3 \qquad u\geq 4m^2, \quad s,t\leq 0, \end{aligned} \tag{40}$$

where s,t are real. **Discuss the Mandelstam plane.**

As we will discuss in much detail below it follows from general principles that the physical amplitude $T(s,t)$ is a boundary value of an analytic function, namely

$$T(s,t) = \lim_{\epsilon\to 0} T(s+i\epsilon,t). \tag{41}$$

Therefore the object of study is an analytic function of two complex variables s and t. Moreover, a single analytic function describes all three different physical regions (40).

3.3 Unitarity for $2\to 2$ scattering

We now discuss the implication of unitarity on the connected $2\to 2$ amplitude, $T(s,t)$. It is common to write the S-matrix as follows

$$\hat{S} \equiv \hat{1} + i\hat{T} \ . \tag{42}$$

From this definition, it follows that

$$S^c_{2,2}(p_1,p_2|p_3,p_4) = i\langle p_3,p_4|\hat{T}|p_2,p_1\rangle = i(2\pi)^d\delta^d\left(p_1+p_2-p_3-p_4\right)T(s,t) \ . \tag{43}$$

[7] In $d=3$ we can also have $\epsilon_{\mu\nu\rho}p_1^\mu p_2^\nu p_3^\rho\ \tilde{T}(s,t)$. We assume that this structure is absent.

In terms of the T-matrix, the unitarity of the S-matrix $\hat{S}\hat{S}^\dagger = \hat{1}$ reads

$$\frac{1}{i}\langle p_3, p_4|\hat{T} - \hat{T}^\dagger|p_2, p_1\rangle = \langle p_3, p_4|\hat{T}\cdot\hat{T}^\dagger|p_2, p_1\rangle \tag{44}$$

$$= \sum_{n=2}^{\infty}\int d\mu_n \langle p_3, p_4|\hat{T}|\{q_i\}_{i=1}^n\rangle\,\langle\{q_i\}_{i=1}^n|\hat{T}^\dagger|p_2, p_1\rangle\ .$$

where in the last step we have inserted a complete basis of states. Due to momentum conservation, an $2n$-particle intermediate state can only contribute if $s > (nm)^2$. Otherwise, there is no enough energy to create an on-shell state with n particles. In particular, for $4m^2 < s < 9m^2$, only two particle states are kinematically allowed. For $9m^2 < s < 16m^2$ two- and three-particle states are allowed in the final state, etc.

Recall that the we adopt the $i\epsilon$-prescription and take

$$\langle p_3, p_4|\hat{T}|p_2, p_1\rangle = \lim_{\epsilon\to 0} T(s + i\epsilon, t). \tag{45}$$

For the conjugate amplitude we get

$$\langle p_3, p_4|\hat{T}^\dagger|p_2, p_1\rangle = \Big(\langle p_2, p_1|\hat{T}|p_3, p_4\rangle\Big)^* = \Big(\langle p_3, p_4|\hat{T}|p_1, p_2\rangle\Big)^*, \tag{46}$$

where in the last line we used the fact that ingoing and outgoing momenta are related by a symmetry transformation (rotation). Applying this relation to (44) below $s < 4m^2$, we conclude that $T(s,t)$ is real below the cut. We can therefore extend the scattering amplitude in the lower half-plane using

$$T(s^*, t) = \Big(T(s,t)\Big)^*, \tag{47}$$

where he we assumed t to be real. This condition is known as *hermitian analyticity* and we use it below to translate unitarity into the statements about the discontinuity of the amplitude.

Let us consider first the elastic unitarity region, $4m^2 < s < 9m^2$ or $4m^2 < s < 16m^2$ if we in addition impose $\mathbb{Z}_2$ symmetry that prohibits even $\to$ odd transitions. In this regime we can replace the sum in (44) by the first term and the equation closes on the $2 \to 2$ transitions only. This is the elastic unitarity regime. Explicitly, we have

$$2T_s(s,t) = \frac{1}{2}\int\frac{d^{d-1}\vec{q}_3}{(2\pi)^{d-1}(2E_{\vec{q}_3})}\int\frac{d^{d-1}\vec{q}_4}{(2\pi)^{d-1}(2E_{\vec{q}_4})}(2\pi)^d\delta^d(p_1 + p_2 - q_3 - q_4)T^{(+)}(s,t')T^{(-)}(s,t'')\ , \tag{48}$$

where we have introduced the notations

$$T^{(\pm)} \equiv \lim_{\epsilon\to 0} T(s \pm i\epsilon, t)\ , \qquad T_s(s,t) \equiv \frac{1}{2i}\Big(T^{(+)}(s,t) - T^{(-)}(s,t)\Big)\ , \tag{49}$$

and $t' = -(\vec{p}_1 - \vec{q}_3)^2$, $t'' = -(\vec{q}_3 - \vec{p}_3)^2$. In writing the RHS we used (47). The factor $\frac{1}{2}$ is $\frac{1}{n!}$ for $n = 2$ due to the fact that we consider identical particles.

We will now reduce the integral to an integration over the two scattering angles by performing all the kinematical integrations explicitly. For that aim, we first go to the center of mass frame where $\vec{q}_3 = -\vec{q}_4 \equiv k\,\hat{n}_3$, where $\hat{n}_3$ is a unit $d-1$ vector and $k = \sqrt{s - 4m^2}/2$. In these variables the elastic unitarity constraint (48) becomes

$$2T_s(s,t) = \frac{k^{d-2}}{(2\pi)^{d-2}(2E_k)^2}\frac{E_k}{2k}\int d^{d-2}\Omega_{\hat{n}_3}\,T^{(+)}(s,t')\,T^{(-)}(s,t'')\ , \tag{50}$$

where $\frac{E_k}{k}$ is the Jacobian coming from the energy conservation delta-function. The integrand only depends on the two scatttering angles

$$z' = \cos\theta' = \frac{\vec{p}_1 \cdot \vec{n}}{|\vec{p}_1|} \qquad \text{and} \qquad z'' = \cos\theta'' = \frac{\vec{p}_3 \cdot \vec{n}}{|\vec{p}_3|} \ . \tag{51}$$

in terms of which we can write the measure as

$$\int d^{d-2}\Omega_{\vec{n}_3} \equiv (4\pi)^{d-2} \int_{-1}^{1} dz' \int_{-1}^{1} dz''\, \mathcal{P}_d(z,z',z'') \qquad \text{where} \qquad z = \cos\theta = \frac{\vec{p}_1 \cdot \vec{p}_3}{|\vec{p}_1||\vec{p}_3|} = 1 + \frac{2t}{s-4m^2} \ , \tag{52}$$

is the cosine of the external scattering angle. One can show that

$$\mathcal{P}_3(z,z',z'') = \frac{1}{4\pi}\sqrt{1-z^2}\,\delta(1-z^2-z'^2-z''^2+2zz'z'') \ , \tag{53}$$

$$\mathcal{P}_{d>3}(z,z',z'') = \frac{1}{(16\pi)^{\frac{d-3}{2}}\Gamma\left(\frac{d-3}{2}\right)} \frac{(1-z^2)^{\frac{4-d}{2}}}{(1-z^2-z'^2-z''^2+2zz'z'')^{\frac{5-d}{2}}}\,\Theta(1-z^2-z'^2-z''^2+2zz'z'') \ .$$

Using that $E_k = \sqrt{s}/2$, we can write (50) covariantly as

$$\boxed{T_s(s,t) = \frac{(s-4m^2)^{\frac{d-3}{2}}}{16\sqrt{s}} \int_{-1}^{1} dz' \int_{-1}^{1} dz''\, \mathcal{P}_d(z,z',z'')\, T^{(+)}(s,t(z'))\, T^{(-)}(s,t(z''))} \ , \tag{54}$$

where $t(x) \equiv -(s-4m^2)(1-x)/2$, not to be confused with the external momentum transfer t.

For $s > 9m^2$ and general t the unitarity constraint involves scattering elements with more than two particles. To get a constraint on the two particle amplitude, we note that $\hat{T}\cdot\hat{T}^\dagger$ on the right hand side of (44) is a positive semi-definite matrix. Hence, for any state Ψ we have that

$$\langle\Psi|\hat{T}|\{q_i\}_{i=1}^n\rangle\,\langle\{q_i\}_{i=1}^n|\hat{T}^\dagger|\Psi\rangle = |\langle\Psi|\hat{T}|\{q_i\}_{i=1}^n\rangle|^2 \geq 0 \ , \tag{55}$$

and hence, if we drop all the contributions with more than two particles we get an inequality for the $2 \to 2$ scattering matrix

$$\frac{1}{i}\int d\mu(p_1,p_1)d\mu(p_3,p_4)\psi(p_1,p_2)\psi^*(p_3,p_4) \times \langle p_3,p_4|\hat{T}-\hat{T}^\dagger|p_2,p_1\rangle \tag{56}$$
$$\geq \int d\mu(p_1,p_1)d\mu(p_3,p_4)\psi(p_1,p_2)\psi^*(p_3,p_4) \times \int d\mu(q_1,q_2)\langle p_3,p_4|\hat{T}|q_1,q_2\rangle\,\langle q_1,q_2|\hat{T}^\dagger|p_2,p_1\rangle \geq 0 \ .$$

For example, if we peak a wave function that consists of two particles with a specific momenta then we have the the amplitude in the forward limit where $p_3 = p_1$ and $p_4 = p_2$. For this choice of wave function, the unitarity constraint (56) becomes

$$T_s(s,0) \geq \frac{(s-4m^2)^{\frac{d-3}{2}}}{16\sqrt{s}} \int_{-1}^{1} dz' \int_{-1}^{1} dz''\, \mathcal{P}_d(1,z',z'')\, T^{(+)}(s,t(z'))\, T^{(-)}(s,t(z'')) \ . \tag{57}$$

3.4 Cross sections

The unitarity relation in the forward limit is nothing but the optical theorem. In d dimensions it takes the form

$$\mathrm{Im}[T(s,0)] = \sqrt{s(s-4m^2)}\,\sigma_{tot}(s) \ , \tag{58}$$

where $\sigma_{tot}(s)$ is the total cross-section, of dimension $[\sigma(s)] = L^{d-2}$.

3.5 Partial wave expansion

The unitarity of the S-matrix implies the non-linear integral relations that the $2 \to 2$ T-matrix has to satisfy, (54) and (56). To simplify these complicated constraints we choose a wave function Ψ that diagonalize the T-matrix and therefore also the integral kernel in (54), (56). This can be done using the Lorentz symmetry of the problem. Namely, we decompose the amplitude $T(s,t)$ in a complete basis of intermediate states which transform in irreducible representations of the $SO(1,d-1)$ symmetry. These representations are characterised by their energy and the little group $SO(d-1)$ angular momentum in the center of mass frame, E and J. For two particle states the $SO(1,d-1)$ quantum numbers are enough to characterize the states and we have

$$\langle p_1, p_2 | p, J, \vec{m}\rangle \propto \delta^d(p - p_1 - p_2) Y^{(d)}_{J,\vec{m}}(\hat{p}_1) \ , \tag{59}$$

where $p^2 = E^2$, $Y^{(d)}_{J,\vec{m}}$ are the d-dimensional spherical harmonics, and the energies dependant prefactor will not be relevant for us.[8] We can now insert a complete basis to these states to decompose the S-matrix element $\langle p_3, p_4|\hat{T}|p_1, p_2\rangle$ in all possible spins. Because the operator $\hat{T}$ is both, translation and $SO(1,d-1)$ invariant, due to the Wigner-Eckart theorem we have that

$$f_J(p^2) \propto \frac{\langle p, J, \vec{m}|\hat{T}|p, J, \vec{m}\rangle}{\langle p, J, \vec{m}|p, J, \vec{m}\rangle} \ , \tag{60}$$

where the convention dependant proportionality factor is independent of the energy and and the angular momentum $\vec{m}$. These functions are the so called partial wave coefficients, in term of which the amplitude takes the form

$$T(s,t) = \sum_J n_J^{(d)} f_J(s) P_J^{(d)}(\cos\theta) \ , \tag{61}$$

where the sum runs over all spins and $n_J^{(d)}$ are convention dependent normalization factors. Here, $P_J^{(d)}(\cos\theta)$ are the partial waves. They represents the angular dependence of the amplitude due to the exchange of all the states with spin J. A simple way of determining these functions is to go to the center of mass frame and act with the $SO(d-1)$ quadratic Casimir on the two outgoing particle, while holding the momentum of the two incoming particles fixed. This equation takes the form

$$\left[(1-z^2)^{\frac{4-d}{2}}\frac{d}{dz}(1-z^2)^{\frac{d-2}{2}}\frac{d}{dz} + J(J+d-3)\right] P_J^{(d)}(z) = 0 \ , \tag{62}$$

[8] For $d=4$ that factor is $\sqrt{\frac{E_{\vec{p}}}{|\vec{p}_1| E_{\vec{p}_1} E_{\vec{p}_2}}}$, see [17] for details.

where $z = \cos\theta$ is cosine of the scattering angle (52). This second order differential equation has two independent solutions. Spin J unitary representations are decomposed from states with angular momentum in the plane of scattering ranging between $-J$ and J. Hence, the corresponding solution of (62) is a degree J polynomyal of $\cos\theta$ that is given by

$$P_J^{(d)}(z) = {}_2F_1\left(-J, J+d-3, \frac{d-2}{2}, \frac{1-z}{2}\right) = \frac{\Gamma(1+J)\Gamma(d-3)}{\Gamma(J+d-3)} C_J^{(\frac{d-3}{2})}(z), \tag{63}$$

where $C_J^{(\frac{d-3}{2})}(z)$ are the standard Gegenbauer polynomials.

The partial wave coefficients can be extracted from the amplitude using the orthogonality relation of these polynomials

$$\frac{1}{2}\int_{-1}^{1} dz\,(1-z^2)^{\frac{d-4}{2}} P_J^{(d)}(z) P_{\tilde{J}}^{(d)}(z) = \frac{\delta_{J\tilde{J}}}{\mathcal{N}_d\, n_J^{(d)}}\ . \tag{64}$$

Here we have chosen the convention

$$\mathcal{N}_d = \frac{(16\pi)^{\frac{2-d}{2}}}{\Gamma\left(\frac{d-2}{2}\right)}\ , \qquad n_J^{(d)} = \frac{(4\pi)^{\frac{d}{2}}(d+2J-3)\Gamma(d+J-3)}{\pi\,\Gamma\left(\frac{d-2}{2}\right)\Gamma(J+1)}\ , \tag{65}$$

for which the unitarity constraint presented below takes a simple form. In this convention we have

$$f_J(s) = \frac{\mathcal{N}_d}{2}\int_{-1}^{1} dz\,(1-z^2)^{\frac{d-4}{2}} P_J^{(d)}(z)\, T\left(s, t(z)\right)\ . \tag{66}$$

Because the S-matrix is diagonal in the spin basis, so does the unitary constraint. We consider first the elastic regime $4m^2 < s < 16m^2$ where this constraints takes the form (54). Using (66), we project both sides to a fixed spin J. On the left hand side to find the imaginary part of the partial wave coefficient, $\text{Im} f_J(s)$. On the right hand side, it is useful to first represent the kernel as a sum over partial waves of z_1, z_2 and z. Because this kernel represents the angular integration in (50), its partial wave decomposition must also be diagonal in spin. It takes that form

$$\mathcal{P}_d(z, z', z'') = \frac{1}{2}\mathcal{N}_d^2 (1-z'^2)^{\frac{d-4}{2}}(1-z''^2)^{\frac{d-4}{2}} \sum_{J=0}^{\infty} n_J^{(d)} P_J^{(d)}(z) P_J^{(d)}(z') P_J^{(d)}(z'')\ . \tag{67}$$

Exercise: Check this relation.

Using (66) for the three integrals, we arrive at the *elastic unitarity* constraint

$$\boxed{2\text{Im} f_J(s) = \frac{(s-4m^2)^{\frac{d-3}{2}}}{\sqrt{s}} |f_J(s)|^2}\ , \tag{68}$$

or equivalently

$$|S_J(s)| = 1\ , \qquad \text{with} \qquad S_J(s) \equiv 1 + i\frac{(s-4m^2)^{\frac{d-3}{2}}}{\sqrt{s}} f_J(s)\ . \tag{69}$$

Here 1 can be traced to back to $\hat{1}$ in (42). In this way the trivial unitary S-matrix $\hat{S} = \hat{\mathbf{1}}$ becomes $S_J = 1$ in the partial wave basis.

The solution to this is

$$f_J(s) = \frac{\sqrt{s}}{(s-4m^2)^{\frac{d-3}{2}}} i(1 - e^{2i\delta_J(s)}) \ , \tag{70}$$

with $\delta_J(s)$ being real for $4m^2 < s < 16m^2$ and is called the scattering phase.

Similarly to the above, for $s > 16m^2$ we chose $\psi(p_1, p_2) = \langle p_1, p_2 | p, J, \vec{m}\rangle$ in (56). In that way we arrive at the same equation, but with an inequality instead of an equality

$$2\text{Im} f_J(s) \geq \frac{(s-4m^2)^{\frac{d-3}{2}}}{\sqrt{s}} |f_J(s)|^2 \ . \tag{71}$$

or, equivalently,

$$|S_J(s)| \leq 1 : \qquad \text{Im}[\delta_J(s)] \geq 0 \ . \tag{72}$$

We close this section with a short discussion on the rage of convergence of the partial wave sum (61) for fixed physical s as a function of $\cos\theta$. As we discuss below, the amplitude $T(s, \cos\theta)$ is analytic inside the small Lehmann-Martin ellipse and its absorptive part, $T_s(s, \cos\theta)$ is analytic inside the large Lehmann-Martin ellipse. These ellipses have foci at $\cos\theta = \pm 1$ and semi-major axis z_{small} and z_{large}. Correspondingly, inside these ellipses the sum (61) and its discontinuity converge.

In the case of scattering of identical lightest particles which is our main interest we have

$$z_{\text{small}} = 1 + \frac{8m^2}{s-4m^2} \ , \qquad z_{\text{large}} = 2z_{\text{small}}^2 - 1 = 1 + \frac{32m^4}{(s-4m^2)^2} \ . \tag{73}$$

In the case of scattering of massless particles, e.g. gravitons, the situations is more complicated and partial wave expansion converges for real angles point-wise in $d > 7$ only. For $5 \leq d \leq 7$ one can use a smeared version of the partial wave expansion.

3.6 Impact parameter representation

To develop some intuition it is useful to connect the formulas above to the so-called *impact parameter representation*. Physically, instead of considering scattering of particles of given momenta we can consider scattering of wave-packets at fixed separation in the transverse space (or fixed impact parameter).

To discuss the impact parameter representation of the amplitude let us consider again the partial wave expansion

$$T(s,t) = \sum_J n_J^{(d)} f_J(s) P_J^{(d)}(1 + \frac{2t}{s-4m^2}) \ . \tag{74}$$

We next consider the following $J \to \infty$, $s \to \infty$ with the ratio

$$b \equiv \frac{2J}{\sqrt{s}} \tag{75}$$

held fixed. We will see below that b plays the role of the impact parameter.

A useful identity to understand this limit is the following

$$\lim_{J\to\infty} P_J^{(d)}(1+\frac{tb^2}{2J^2}) = 2^{\frac{d}{2}-2}\Gamma(\frac{d}{2}-1)(\sqrt{-t}b)^{2-\frac{d}{2}} J_{\frac{d-4}{2}}(\sqrt{-t}b) + \mathcal{O}(\frac{1}{J}), \tag{76}$$

where $J_\alpha(x)$ is the Bessel function of the first kind.

Switching from the sum $\sum_J \to \frac{\sqrt{s}}{2}\int db$ we get the following representation for the amplitude

$$T(s,t) = i\ 4s \int_0^\infty dbb^{d-3}\left((2\pi)^{\frac{d}{2}-1}(\sqrt{-t}b)^{2-\frac{d}{2}} J_{\frac{d-4}{2}}(\sqrt{-t}b)\right)\left(1-e^{2i\delta(s,b)}\right) + ..., \tag{77}$$

where ... represents terms that are suppressed in the large s limit.

In writing the formula above we assumed that the phase shift $\delta_J(s)$ that enters into the definition of partial waves admit the following limit

$$\lim_{s\to\infty} \delta_{J=\frac{b\sqrt{s}}{2}}(s) \equiv \delta(s,b). \tag{78}$$

The formula (77) can be rewritten using the following identity

$$\int_{-\infty}^{\infty} d^{d-2}\vec{b} e^{i\vec{q}\cdot\vec{b}} f(|\vec{b}|) = (2\pi)^{\frac{d}{2}-1}\int_0^\infty dbb^{d-3}\,(qb)^{2-\frac{d}{2}}\, J_{\frac{d-4}{2}}(qb) f(b). \tag{79}$$

In this way we finally get

$$T(s,-\vec{q}^2) = i\ 4s\int_{-\infty}^{\infty} d^{d-2}\vec{b} e^{i\vec{q}\cdot\vec{b}}\left(1-e^{2i\delta(s,b)}\right) + ...\ . \tag{80}$$

Or, equivalently inverting this relation we get the following expression for the phase shift

$$S(s,b) \equiv 1-e^{2i\delta(s,b)} = -\frac{i}{4s}\int \frac{d^{d-2}\vec{q}}{(2\pi)^{d-2}} e^{-i\vec{q}\cdot\vec{b}} T(s,-\vec{q}^2) + ...\ . \tag{81}$$

This makes the physical meaning of the phase shift manifest. Indeed, to go from the amplitude to the phase shift we perform the Fourier transform in the transferred momentum which fixes the separation between the scattering objects to be $\vec{b}$.

Exercise: Check the derivation of the formulas above.

Unitarity of the partial waves $|S_J(s)| \leq 1$ becomes

$$|S(s,b)| \leq 1, \quad s > 0, \tag{82}$$

which should hold up to terms that decay at large s. This is the condition we encountered in figure 3.

Exercise: (The Froissart bound). Let us consider scattering in a gapped theory (for example scattering of pions in QCD). Imagine we are interested in the behavior of the total cross section (58) at large energies $\frac{s}{m_\pi^2} \gg 1$.

Using (80) we get

$$\sigma_{tot}(s) = \frac{8\pi^{\frac{d}{2}-1}}{\Gamma(\frac{d}{2}-1)}\int_0^\infty db\ b^{d-3}\text{Re}[1-S(s,b)] \tag{83}$$

Since the theory is gapped we expect that at fixed s and large impact parameters the phase shift behaves as $\delta(s,b) \sim s^N e^{-2m_\pi b}$ (this is nothing but the Yukawa potential for the two-pion exchange),[9] where we assume that $N > 0$.

We expect thus that the large impact parameters are suppressed and we can restrict the relevant impact parameters to the region $s^N e^{-2m_\pi b} \sim 1$, or $b \leq b_{\max} = \frac{N}{2m_\pi} \log \frac{s}{s_0}$. In this way we get

$$\sigma_{tot}(s) \leq C \int_0^{b_{\max}} db b^{d-3} \sim C b_{\max}^{d-2} \sim \left(\frac{N}{2m_\pi} \log \frac{s}{s_0} \right)^{d-2} . \tag{84}$$

This is the famous Froissart bound. Below we will consider a much neater and more precise derivation of the same result without any squiggly lines, but the basic result is the same: in a gapped theory the size of a relativistic object grows at most like $\log s$. For a recent discussion of the status of the Froissart bound in QCD at the current accelerator energies see [36] and references therein. The punchline is that there is no reason to expect that we will observe the Froissart bound saturation at the LHC energies.

Exercise: (The Gravitational Froissart bound). Let us consider scattering in a gravitational theory. Imagine we are interested in the behavior of the total cross section (58) at large energies $\frac{s}{m_{Pl}^2} \gg 1$. Derive the expected behavior based on the fact that at large impact parameters gravity becomes classical and the phase shift is thus given $\delta(s,b) = \frac{G_N s}{b^{d-4}}$. Hint: use unitarity at small impact parameters and the explicit expression for $\delta(s,b)$ at large impact parameters.

3.7 Functions versus distributions

Based on the standard QFT axioms, scattering amplitudes, cross sections and partial waves are tempered distributions rather than continuous functions. This leads to various subtleties, sometimes known as *pathologies a-la Martin* [26]. For example, we can imagine total cross-section having singularities which are local in energy variable s, that are not detectable by the finite resolution experiments. As such it is hard to exclude them based on physical grounds. One way to produce such a singularity is to consider an infinite number of resonances that accumulate on the real axis, see [26].

To the best of our knowledge there is no known, first principle argument that can exclude these possibilities. One way to eliminate them is to simply assume that various cross sections are actually real analytic functions of energy s rather than distributions. We add this to the list of our assumptions. It would be very interesting to see if it is possible to establish this rigorously.

4 Analyticity

Analyticity is the central element of S-matrix theory. It is a nontrivial consequence of the basic principles of QFT and it is believed to hold more generally, e.g. in gravitational theories. It also measures our power to derive results in S-matrix theory: the more analyticity we have, the better results we can derive.

It is natural to divide the discussion of analyticity into several parts. The first part of the discussion is the so-called axiomatic analyticity: the domain of analyticity that can be derived from QFT axioms supplied by the LSZ reduction formula and the technique of analytic completion. The

[9] Two-pion exchange is the minimal mass of the exchanged state because pions are pseudoscalars.

most famous outcome of this analysis is the Froissart bound on the total cross-section that we have already briefly discussed above and derive more carefully below.

The second part describes the analytic structure of the amplitudes that is consistent with unitarity. By analytic continuation of unitarity relations we naturally arrive to the notion of the Landau equations and Landau curves. These are best understood in perturbation theory even though the analytic structure derived is nonperturbative. One famous outcome of this analysis is a failed attempt to prove the so-called Mandelstam representation with many exciting singularities discovered on the way (anomalous thresholds, cusps, acnodes, etc).

Another outcome of this analysis is the idea of *maximal analyticity*: the maximal domain of analyticity that is consistent with unitarity. This is something that has not been rigorously proven but surves as a working hypothesis in many papers on the S-matrix bootstrap.

4.1 Axiomatic analyticity: rigorous domain of analyticity derived from QFT axioms

For a very nice and brief compendium of results derived using the axiomatic QFT methods, see [4]. We refer to this review and the original references listed there for details. Another source of detailed derivations is the review by Sommer [5].

The purpose of this section is to briefly review the relevant techniques of establishing analytic properties of the scattering amplitudes starting from basic axioms of massive QFT and techniques to analytically continue functions of two complex variables:

1. Lorentz invariance.
2. Causality.
3. Mass gap and spectrality.
4. Analytic completion.
5. Temperedness.
6. Extra technical assumptions?

Deriving analytic properties of scattering amplitudes proceed in the following steps:

1. Express scattering amplitude in terms of the correlation function using the LSZ-type reduction formulas.
2. Using causality or commutativity of operators at space-like separated points to establish analyticity in the upper (lower) half-plane in a correspondent complex variable.
3. Use completeness of asymptotic states to relate different regions of analyticity. In this way one establishes analyticity in the so-called primitive domain.
4. Use techniques of analytic completion to extend the primitive analyticity domain.
5. Go on-shell and establish analyticity of the on-shell scattering amplitudes.
6. Extend the region of analyticity using unitarity.

An interesting fact worth emphasizing is that skipping step 4 of analytic completion leads to empty region of analyticity of scattering amplitudes when going on-shell.

Let us proceed in order of the strength of the results derived from the axioms. The starting point is the statement the result that the scattering amplitude is polynomially bounded

$$|T(s,t)| < |s|^N, \quad t < 0. \tag{85}$$

Moreover for fixed $t < 0$ the amplitude is analytic in the complex s-plane except for cuts related to unitarity. Analyticity and sub-exponential growth in the upper half-plane follow from causality, and the polynomial boundedness eventually follow from temperedness. This simple structure of the s-complex plane requires some conditions on the masses of the scattered particles. For example it holds for the scattering of lightest particles in the theory. For generally, there is a region $|s| < R(t)$ that can contain "dragons", such as anomalous thresholds. Existence of (85) together with the associated analyticity structure leads to dispersion relations to be discussed below.

For elastic scattering, one can similarly show that for real and fixed s the amplitude is analytic inside *the Lehmann ellipse*.[10] [11]

This is an ellipse with foci at ± 1 and whose size depends on the details of the theory and s. Analyticity inside the Lehmann ellipse extends analyticity in t to some positive values $t_L(s)$. The most important property of the Lehmann ellipse to keep in mind is that it shrinks with energy

$$\text{The Lehmann ellipse}: \quad t_L(s) \sim \frac{C}{s}, \quad s \to \infty. \tag{87}$$

One of the most famous results in S-matrix theory is the work by AndrÃ© Martin [3] where he used unitarity to extend analyticity in t beyond the Lehmann ellipse. His starting point are dispersion relations with N subtractions and the outcome is the statement that $T(s,t)$ is *analytic* and *polynomially bounded* in s for $|t| \leq t_0$.

$$\text{The Martin's extension}: \quad |t| \leq t_0, \tag{88}$$

where t_0 *does not* depend on s. For the scattering of identical particles $t_0 = 4m^2$.[12]

At this point we are ready to derive the most famous result of S-matrix theory, namely the Froissart-Martin bound.

[10] Why are we talking about ellipses? This is related to convergence of the partial wave expansion, which can be seen using Neumann's argument [7]. Consider a function $f(z)$ analytic inside some region $\mathcal{C}$ which includes the $[-1,1]$ interval. We can then write

$$f(z) = \oint_\gamma \frac{dt}{2\pi i}\frac{f(t)}{t-z} = \oint_\gamma \frac{dt}{2\pi i}\sum_{J=0}^{\infty} n_J^{(d)} P_J^{(d)}(z)\left(\mathcal{N}_d (t^2-1)^{\frac{d-4}{2}} Q_J^{(d)}(t) f(t)\right) = \sum_{J=0}^{\infty} n_J^{(d)} f_J P_J^{(d)}(z), \tag{86}$$

where $\gamma \in \mathcal{C}$ is some contour that wraps the interval $[-1,1]$ counterclockwise and contains z inside the integration contour. $Q_J^{(d)}(t)$ are known functions that we will encounter in our lectures soon. To exchange the summation and integration, we also used that given z, $\frac{1}{t-z} = \mathcal{N}_d(t^2-1)^{\frac{d-4}{2}} \sum_{J=0}^{\infty} n_J^{(d)} P_J^{(d)}(z) Q_J^{(d)}(t)$ converges uniformly in t as long as t is outside the ellipse with foci at -1 and 1 that passes through z. We also used the relation between $Q_J^{(d)}(z)$ and $P_J^{(d)}(z)$ which will be explained below.

[11] There is a story related to this [44]: "It was in the middle of 1950th at the Conference held in Dubna, where Lehmann reported his above mentioned result in the first time. The attended there Landau has argued during the talk that the Lehmann result was wrong because in his opinion two-body elastic scattering amplitude must be analytic function in the whole complex plane. Harry Lehmann wittily replied: Landau is a big man, so he needs analyticity of the amplitude in the whole complex plane, but I am a small man, and analyticity of the amplitude in the ellipse is enough for me."

[12] Extending it any further would have been impossible since the two-particle cut starts at $t = 4m^2$. In this sense Martin's extension is in the spirit of maximal analyticity to be discussed below.

4.2 Derivation of the Froissart-Martin bound

For a careful discussion of various subtleties that go into the argument check [6].

In the previous section we had a qualitative derivation of the Froissart bound under some mild assumption. We are now ready to prove the precise bound. Let us start with the quantity that we would like to bound, namely the averaged cross section

$$\bar{\sigma}_{\text{tot}}(s) = \frac{1}{s-4m^2}\int_{4m^2}^{s} ds'(s'-4m^2)\sigma_{\text{tot}}(s'). \tag{89}$$

Considering things on average is a more careful way of discussing the bound on cross-section which does not require an assumption about point-like behavior of $\sigma_{\text{tot}}(s)$, see section 3.7. For simplicity we restrict the consideration to $d = 4$.

As a first step, we plug the partial wave expansion for $\sigma_{\text{tot}}(s)$ inside the integral (89) and we split the sum into low spins and high spins. We get using (61)

$$\sigma_{\text{tot}}(s) = \frac{1}{\sqrt{s(s-4m^2)}}\text{Im}[T(s,0)] = \frac{1}{\sqrt{s(s-4m^2)}}\sum_J n_J^{(d=4)}\text{Im}f_J(s),$$

$$\bar{\sigma}_{\text{tot}}(s) = \frac{1}{s-4m^2}\Big(\sum_{J=0}^{L}\int_{4m^2}^{s} ds'\sqrt{\frac{s'-4m^2}{s'}}n_J^{(d=4)}\text{Im}f_J(s') + \sum_{J=L+2}^{\infty}\int_{4m^2}^{s} ds'\sqrt{\frac{s'-4m^2}{s'}}n_J^{(d=4)}\text{Im}f_J(s')\Big). \tag{90}$$

In the first term we use unitarity (70) to get

$$0 \le \text{Im}f_J(s) = \frac{\sqrt{s}}{(s-4m^2)^{\frac{d-3}{2}}}\underbrace{\text{Re}[1-e^{2i\delta_J(s)}]}_{\text{Im}\delta_J(s)\ge 0} \le \frac{2\sqrt{s}}{(s-4m^2)^{\frac{d-3}{2}}}. \tag{91}$$

In this way we get (using $n_J^{(d=4)} = 16\pi(2J+1)$)

$$\begin{aligned}&\frac{16\pi}{s-4m^2}\sum_{J=0}^{L}(2J+1)\int_{4m^2}^{s} ds'\sqrt{\frac{s'-4m^2}{s'}}\text{Im}f_J(s')\\ &\le \frac{32\pi}{s-4m^2}\sum_{J=0}^{L}(2J+1)\int_{4m^2}^{s} ds' = 16\pi(L+1)(L+2),\end{aligned} \tag{92}$$

where the sum goes over even J only (because the particles are identical) and L is even as well. At this point we simply used unitarity.

Our next task is to bound the sum over the high spin partial waves $\sum_{J=L+2}^{\infty}$ in (90). It is at this point that we use Martin's extension of analyticity and polynomial boundedness. For simplicity we set $t = t_0 = 4m^2$ but any positive t_0 would do. *Polynomial boundedness* states that there exists N such that

$$a_N \equiv \int_{4m^2}^{\infty}\frac{ds'}{(s')^{N+1}}\text{Im}[T(s',4m^2)] < \infty. \tag{93}$$

In four dimensions $P_J^{(d=4)}(z)$ are the usual Legendre polynomials $P_J(z)$, where $z = 1+2\frac{t}{s-4m^2}$. It is straightforward to check that for $z > 1$ (which corresponds to $s \ge 4m^2$ and $t > 0$) they are positive and monotonically increasing functions of z and J. Again we plug the partial wave expansion into (93) to get

$$
\begin{aligned}
a_N &= 16\pi \int_{4m^2}^{\infty} \frac{ds'}{(s')^{N+1}} \sum_{J=0}^{\infty} (2J+1) \mathrm{Im} f_J(s') P_J(1+\frac{8m^2}{s'-4m^2}) \\
&\geq 16\pi \int_{4m^2}^{s} \frac{ds'}{(s')^{N+1}} \sum_{J=L+2}^{\infty} (2J+1) \mathrm{Im} f_J(s') P_J(1+\frac{8m^2}{s'-4m^2}) \\
&\geq 16\pi \frac{P_{L+2}(1+\frac{8m^2}{s-4m^2})}{s^{N+1}} \sqrt{\frac{s}{s-4m^2}} \int_{4m^2}^{s} ds' \sum_{J=L+2}^{\infty} (2J+1) \sqrt{\frac{s'-4m^2}{s'}} \mathrm{Im} f_J(s'). \qquad (94)
\end{aligned}
$$

In this way we have bounded the higher spin sum in (90) as follows

$$
\sum_{J=L+2}^{\infty} \int_{4m^2}^{s} ds' \sqrt{\frac{s'-4m^2}{s'}} n_J^{(d=4)} \mathrm{Im} f_J(s') \leq \frac{a_N}{\frac{P_{L+2}(1+\frac{8m^2}{s-4m^2})}{s^{N+1}} \sqrt{\frac{s}{s-4m^2}}}. \qquad (95)
$$

We can thus write the bound on the total cross section as follows

$$
\bar{\sigma}_{\text{tot}}(s) \leq 16\pi (L+1)(L+2) + \sqrt{\frac{s-4m^2}{s}} \frac{a_N s^{N+1}}{P_{L+2}(1+\frac{8m^2}{s-4m^2})}. \qquad (96)
$$

Next we use the following property of the Legendre polynomials

$$
P_J(z) \geq \frac{\phi_0}{\pi} \Big(z + \cos\phi_0 \sqrt{z^2-1} \Big)^J, \quad z \geq 1, \quad 0 < \phi_0 < \pi. \qquad (97)
$$

To simplify the expressions we now take $L \gg 1$ and $s \gg 4m^2$. We then get

$$
\sqrt{\frac{s-4m^2}{s}} \frac{a_N s^{N+1}}{P_{L+2}(1+\frac{8m^2}{s-4m^2})} \simeq \frac{a_N s^{N+1}}{P_{L+2}(1+\frac{8m^2}{s})} \lesssim \frac{\pi a_N s^{N+1}}{\phi_0} \Big(1 + \frac{4m}{\sqrt{s}} \cos\phi_0 \Big)^{-L-2}. \qquad (98)
$$

We thus get to leading order in s and L

$$
\bar{\sigma}_{\text{tot}}(s) \lesssim 16\pi L^2 + \frac{16\pi^2 a_N s^N}{\phi_0} \Big(1 + \frac{4m}{\sqrt{s}} \cos\phi_0 \Big)^{-L}. \qquad (99)
$$

To derive the final bound we would like to pick L in the expression above to optimize the bound. The optimal choice turns out to be

$$
L_* = \frac{(N-1)\sqrt{s} \log\frac{s}{s_0}}{4m \cos\phi_0} \ . \qquad (100)
$$

Exercise: Derive (100) by extremizing (99) with respect to L. Check that on the solution the second term in (99) goes to zero as $s \to \infty$.

In this way we finally get

$$
\bar{\sigma}_{\text{tot}}(s) \lesssim 16\pi s \Big(\frac{(N-1) \log\frac{s}{s_0}}{4m \cos\phi_0} \Big)^2. \qquad (101)
$$

This is almost the desired Froissart-Martin bound!

There are still a few wrinkles in the bound above: it depends on ϕ_0, N and s_0. With some extra work, namely by using the result (101) to show that $N = 2$, and setting $\phi_0 = 0$ we get[13]

$$\bar{\sigma}_{\rm tot}(s) \lesssim \frac{\pi}{m^2} s (\log \frac{s}{s_0})^2, \quad s \gg m^2 \tag{102}$$

ABSOLUTE BOUND ON CROSS-SECTIONS AT ALL ENERGIES AND WITHOUT UNKNOWN CONSTANTS

F. J. YNDURAIN
CERN, Geneva, Switzerland

Received 6 February 1970

We prove that the total cross-sections for $\pi\pi$ scattering are bounded at all energies by

$$\frac{1}{s-4\mu^2}\int_{4\mu^2}^{s} ds'(s'-4\mu^2)\sigma_{\rm tot}(s') < \pi\left(\frac{n+1}{n}\right)^2 \frac{s}{\mu^2}\log^2\frac{s}{\mu^2} + 8\pi\left(\frac{n+1}{n}\right)\sqrt{\frac{s}{\mu^2}}\log\frac{s}{\mu^2} + (1280)\pi\mu^3\frac{2^n(n!)^2[2n/(n+1)]^n}{(2n)!}a_2^t s,$$

valid for all $n \geqslant 1$, and where a_2^t is the t channel D wave scattering length. At high energies, the right-hand side becomes essentially the Froissart bound without unknown constants, viz.,

$$\pi[(n+1)/n\mu]^2 s \log^2 s/\mu^2 .$$

Figure 2: The Froissart-Martin bound made precise.

Finally, s_0 can be as well removed from the formula to get a rigorous finite energy bound in terms of parameters that are measurable in the experiment only, see [9] and figure 2.

A quick recap:

Let us recapitulate again the two main results of this heroic exercises for the scattering of the identical (lightest to avoid dragons) massive particles. For $|t| < 4m^2$ the amplitude is analytic in the cut s-plane. In addition the following results hold:

$$\text{The Froissart} - \text{Martin bound}: \quad |T(s, t=0)| < 2\pi \frac{s}{m^2}(\log \frac{s}{m^2})^2, \tag{103}$$

$$\text{Two subtractions}: \quad \lim_{|s|\to\infty} \frac{|T(s,t)|}{s^2} = 0, \quad |t| < 4m^2, \tag{104}$$

where the last fact has been proven in [11].

Similarly, using the fact that

$$|P_J(\cos\theta)| \leq 2\sqrt{\frac{1}{\pi(2J+1)\sin\theta}}, \quad 0 < \theta < \pi, \tag{105}$$

[13]In the original paper [8] Froissart used the Mandelstam representation (to be discussed below) which does not follow from axiomatic analyticity and unitarity. So it does make sense to add Martin to its name.

and the bounds above one gets

$$\text{The fixed} - \text{angle bound}: \quad |T(s, t(\cos\theta))| < C\frac{s^{3/4}(\log s)^{3/2}}{(\sin\theta)^{1/2}}. \tag{106}$$

4.3 Crossing

Let us consider first the scattering of identical scalar particles (let us also assume that these are the lightest particles in the theory, e.g. pions in QCD). In this case all three channels (40) describe the same physical process and are thus trivially equal. Moreover in each channel the relevant amplitude is a boundary value obtained by taking the limit from the region of analyticity (45).

What is nontrivial is that in fact all these channels can be described by *a single analytic function* $T(s,t)$ that satisfies

$$T(s,t) = T(t,s) = T(u,t), \quad s + t + u = 4m^2, \tag{107}$$

and describes different scattering channels which are connected through the region of analyticity. This property is called *crossing*. We thus have a single function that encodes different channels.

The situation is simple in the case of scattering of the lightest particles in the theory. In this case the amplitude is analytic and real inside *the Mandelstam triangle*

$$\text{Mandelstam triangle}: \quad 0 < s, t, u < 4m^2. \tag{108}$$

This is unphysical region below all the cuts. Now by analytically continue through the Mandelstam triangle we can connect any two channels to each other.

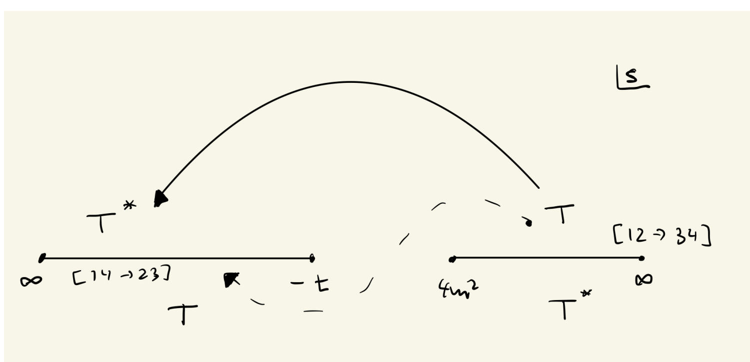

Figure 3: The s-plane without anomalous thresholds.

Exercise: Check that as an immediate consequence of crossing the partial waves $f_J(s)$ as well as the phase shift $\delta(s,b)$ have *the left-cut* starting at $s = 0$. This is what was depicted in figure 3

For scattering of particles of general masses the situation is more complicated. In this case we do not know the analyticity properties of the amplitude in some region close to the origin (there are *anomalous thresholds* to be discussed later). Therefore we cannot go from the s- to the u- channel

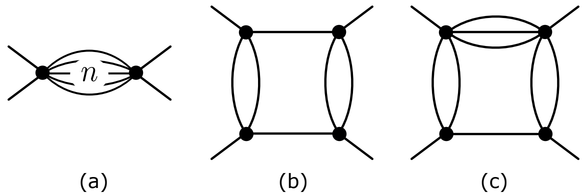

Figure 4: A few simplest examples of graphs that represent various singularities of the $2 \to 2$ scattering amplitude. a) The bubble diagram represents multi-particle normal thresholds. b) The two-particle box diagram. It represents a Landau curve along which the scattering amplitude develops double discontinuity. c) The four-particle box diagram. This diagram corresponds to four-particle scattering both in the s- and in the t-channel. In this paper we systematically study the graphs of this type and the corresponding Landau curves.

directly. Luckily, in this case there is the so-called Bros-Epstein-Glaser path which still allows to connect different channels via analytic continuation. The basic idea is that we first perform $s \leftrightarrow t$ crossing by keeping u fixed and analytically continuing the amplitude through the upper half-plane. In this way we connect $T_{12\to 34}$ to $T^*_{13\to 24}$. We then do the continuation $t \leftrightarrow u$ with s fixed which maps $T^*_{13\to 24}$ to $T_{23\to 14} = T_{14\to 23}$.[14]

In the arguments above we assumed that all the particles are identical and that they coincide with their anti-particle (same charge, opposite quantum numbers). More generally, the transformation of crossing relates amplitudes $i, X \to j, Y$ and $\bar{j}, X \to \bar{i}, Y$. By performing the crossing transformation twice we also obtain the statement known as the CPT theorem, namely that

$$T_{a,b\to c,d}(s,t) = T_{\bar{c},\bar{d}\to \bar{a},\bar{b}}(s,t). \tag{109}$$

In this case the first crossing transformation relates $T_{a,b\to c,d}$ to $T^*_{a,\bar{c}\to \bar{b},d}$, and the second one maps $T^*_{a,\bar{c}\to\bar{b},d}$ to $T_{\bar{c},\bar{d}\to\bar{a},\bar{b}}(s,t)$. Similarly, given the s-channel process $T_{1,2\to 3,4}(s,t)$ $(s > (m_1+m_2)^2)$, the analytic continuation of $T_{1,2\to 3,4}(s,t)$ to $t > (m_1+m_3)^2$ describes the process $1 + \bar{3} \to \bar{2} + 4$, and similarly for the u-channel.

Crossing symmetry for $2 \to 2$ scattering amplitudes was famously established starting from QFT axioms in the paper by Bros, Epstein, and Glaser [12] (Bros et al also generalized it later to $2 \to 3$). This year Sebastian Mizera gave a general argument for crossing in the planar limit [28]. In particular, see appendix A in that paper regarding subtleties in establishing crossing starting from QFT axioms. See also the lectures by Simon Caron-Huot for a nice discussion of crossing.

4.4 Analyticity from Unitarity

Introduce the relation between analyticity and unitarity. This discussion is taken from [27].

The $2 \to 2$ scattering process is characterized by an analytic function $T(s,t)$ that depends on two independent (complex) Mandelstam variables $s = -(p_1+p_2)^2$ and $t = -(p_1+p_4)^2$, where p_i^μ are the on-shell momenta, $p_i^2 = -m^2$, of the scattered scalar particles.[15]

We would like to understand the minimal set of singularities possessed by $T(s,t)$ as a consequence of unitarity and crossing. Here we aim at revealing an infinite subset of singularities

[14]This is particularly important for crossing for scattering amplitudes of massless particles. In this case, there is never an analog of (108).

[15]The results derived in this paper should equally apply to spinning particles.

associated with multi-particle unitarity. The simplest singularities of this kind are *normal thresholds*. These are branch-point singularities at $s, t, u = (nm)^2$, with $n \geq 2$. Their presence follows directly from unitarity

$$\text{Disc}_s T(s,t) \equiv \frac{T(s+i\epsilon,t) - T(s-i\epsilon,t)}{2i} = \sum\!\!\!\!\!\!\int_n T_{2\to n} T^\dagger_{2\to n}\,,$$
$$\text{with} \quad s \geq 4m^2\,, \quad 4m^2 - s < t < 0\,, \tag{110}$$

and the fact that $T_{2\to n} = 0$ for $s < (nm)^2$. Here, the integral is over the n-particle phase space. To each term in the sum in (110) we can assign a graph. The vertices in this graph represent the amplitudes $T_{2\to n}$, $T^\dagger_{n\to 2}$ and the lines between them represent the n-particle state.

As we analytically continue (110) to $t > 0$, we may encounter discontinuities of $\text{Disc}_s T(s,t)$ in t. For example, consider the term in (110) with $n = 2$. Both $T_{2\to 2}(s,t')$ and $T^\dagger_{2\to 2}(s,t'')$ have a normal 2-particle threshold in the t-channel. These start to contribute to the corresponding phase space integral in (110) at a new branch-point that is located at

$$(s - 4m^2)(t - 16m^2) - 64m^4 = 0\ , \tag{111}$$

along which the scattering amplitude develops double discontinuity, see [35] or the discussion of the box diagram below for details. These curves along which the double discontinuity is developed are called *Landau curves*.

We can assign to this double discontinuity the graph in figure 4.b, where again, the lines represent (on-shell) particles and the vertices represent four-point amplitudes that have been analytically continued outside the regime of real scattering angles.

As we take $s > 16m^2$ more n's contribute to (110) and more singularities are produced by the corresponding phase space integration. For example, the integration over the four-particle phase space ($n = 4$) can produce a cut of $\text{Disc}_s T(s,t)$ in t that results from the analytically continued two-particle normal threshold of $T_{2\to 4}$ and $T^\dagger_{2\to 4}$. The graph that represents this contribution to the double discontinuity $\text{Disc}_t\text{Disc}_s T(s,t)$ is plotted in figure 4.c.

Similarly, for any singularity that follows from multiple iteration of (analytically continued) unitarity we can associate a corresponding graph. By iteration of unitarity we mean the double discontinuity of the amplitude that is generated from a singularity of $T_{2\to n}$ and another singularity of $T^\dagger_{2\to n}$, through the analytic continuation of the phase space integration in (110) to $t > 0$. The singularities of $T_{2\to n}$ and $T^\dagger_{2\to n}$ themselves follows from analytically continued unitarity in a similar fashion. The graph that we associate to such a contribution to $\text{Disc}_t\text{Disc}_s T(s,t)$ is defined recursively, by gluing together a graph that represents a singularity of $T_{2\to n}$ with a one that represents a singularity $T^\dagger_{2\to n}$ with n-lines.

To enumerate all singularities that emerge in this way, we can go in the opposite direction and first enumerate all graphs that may result in a singularity of the amplitude. Whether a given graph leads to a singularity of the amplitude in a certain region in the complex s, t planes is a kinematical question that does not depend on the details of the sub-amplitudes, represented by the vertices in the graph.[16] Hence, to answer this question we can equivalently take them to be constants.

[16]Note that the described way of generating new singularities from old ones involves analytic continuation of the amplitudes. It might happen that due to some special properties of the amplitude, the expected singularity is not there. Here we assume that this does not happens and expect the singularities which follow from unitarity to be generically present.

After doing so, it becomes evident that the same singularity, if it exists, is also generated by the Feynman diagram that coincides with the graph obtained from unitarity. The relevant singularity of the diagram comes from the region of loop integration where all propagators go on-shell. Other singularities of Feynman diagrams may result from a region of the loop integration where only a subset of propagators is on-shell. Those propagators that remain off-shell at the locus of a given singularity can thus be regarded as part of an higher point vertex that is not constant. For example, the Feynman diagrams that correspond to the graph in figure 4.a with two lines and the graph in figure 4.b, both have normal threshold at $s = 4m^2$. Hence, the set of all singularities of a Feynmann diagram includes the singularities of the corresponding graph and graphs obtained from it by collapsing some subset of lines into vertices with more legs. This operation is called a contraction.

If a generic diagram has an n-particle cut then it has a normal threshold starting at n^2m^2 in s, t or u (depending on which external legs are considered incoming/outgoing). This can be seen by contracting the rest of the lines into a bubble diagram as in figure 4.a, with n legs.

In this way we immediately conclude that figure 4.b has normal thresholds at $s = 4m^2$, $t = 16m^2$, and figure 4.c at $s = 16m^2$, $t = 16m^2$.

In conclusion, to enumerate the singularities that follow from unitarity we can equally enumerate the singularities of Feynman diagrams.

In this classification, the Feynman diagrams are only used as a tool to study the location of kinematical singularities of a non-perturbative amplitude. For more than two intermediate particles, we find this tool more practical than directly analyzing the analytic continuation of the unitarity relation (110).

The locations of singularities of Feynmann diagrams can be found using the Landau equations. These are summarized below and we refer the reader, for example, to [2, 28] for a detailed review.

A *leading singularity* of a Feynmann diagram is a singularity that coincides with that of the corresponding graph.

Therefore we may restrict our dissection to singularities of this type only. The Landau equations that correspond to such a singularity are

1. All propagators are on-shell, $k_i^2 = m_i^2$, where the index $i = 1, \ldots, P$ labels all the propagators and k_i's are oriented momenta that flow through them.

2. At each vertex v, the momentum is conserved, $\sum_{j\in v} \pm k_i^\mu = 0$, with + (−) for ingoing (outgoing) momenta.

3. For any loop l, the momenta satisfy
   $\sum_{j\in l} \pm\alpha_j k_i^\mu = 0$, with + (−) sign for momenta along (opposite) the orientation of the loop, and non-zero coefficients, $\alpha_i \neq 0$.

Two solutions that are related by an overall rescaling of the coefficients corresponds to the same singularity. We may therefore normalize them such that $\sum \alpha_i = 1$.

For any solution to these equations we can associate a story in complexified spacetime. In this story the Feynman parameters, α_i, are the proper times of on-shell particles, $k_i^2 = m_i^2$, that propagate along the spacetime interval $\Delta x_i^\mu = \alpha_i k_i^\mu$. Every vertex represents a scattering of these particles that takes place at a point. The spacetime interval between two vertices should not depend on the path between the vertices. This means that for a closed path (i.e. a loop) we have $\sum_{i\in l} \Delta x_i^\mu = 0$.

No general answer is known to the question of which parts of the Landau curve lead to singularities on the physical sheet (which is our main interest here).

4.5 Anomalous threshold: the triangle graph

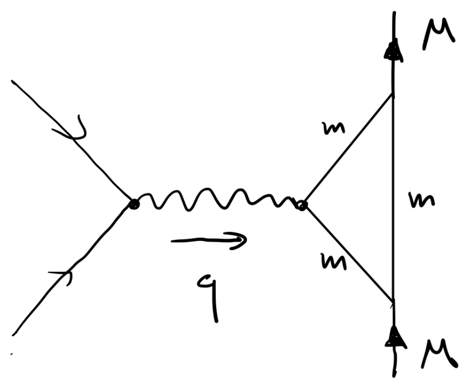

Figure 5: Electromagnetic form factor. We can think of a neutral particle of mass M that couples to an electron-positron pair which is then probed by a virtual photon (say, emitted by another electron).

How does a photon see something, for example, a positronium or an electron? It does it via an electromagnetic form factor

$$F(q^2) \equiv \langle p+q|J_{EM}(q)|p\rangle. \tag{112}$$

In partcular, taking $q^2 = -\vec{q}^2$ and doing the Fourier transform we get[17]

$$\tilde{F}(r = |\vec{x}|) = \int d^{d-1}\vec{x} e^{i\vec{q}\cdot\vec{x}} F(\vec{q}^2). \tag{113}$$

for the effective distribution of charge in the object. If the object is point-like, e.g. an electron, we will get to leading-order a $\delta(r)$. On the other hand, if the object has some electromagnetic size r_0 we expect the form factor to be $\tilde{F}(r) \sim e^{-\frac{r}{r_0}}$.

The size of the object is controlled by the closest to the real axis non-analyticity in $\vec{q}^2$. For example, taking $F(\vec{q}^2) \sim \frac{1}{\vec{q}^2+m^2}$ we will get the Yukawa potental $F(r) \sim e^{-mr}$ coming from the pole $\vec{q}^2 = -m^2$. Consider next a slightly more complicated triangle diagram, see figure 5. From the point of view of this diagram the size of the object is controlled by the leading singularity in $s = -\vec{q}^2$. For example, for the analytic structure presented in figure 3 we will get that the size of the form factor is $F(r) \sim e^{-2mr}$ coming from the fact that the nearest singularity is at $s = 4m^2$.

Let us now compute the triangle graph directly. For simplicity we imagine that all the propagators are scalar (this will not affect the analytic structure of the diagram). We also set $d = 4$. In

[17] At this point, we randomly have switched to the mostly minus signature.

this way we get

$$F(q^2, M^2) = \int_0^1 \prod_{i=1}^{3} d\alpha_i \frac{\delta(1 - \sum_{i=1}^{3} \alpha_i)}{m^2 - (\alpha_1 + \alpha_2)\alpha_3 M^2 - \alpha_1 \alpha_2 q^2}. \tag{114}$$

The result for this integral takes the form (to simplify the result we differentiate with respect to m^2)

$$\partial_{m^2} F(q^2, M^2) = \frac{\frac{(2M^2 - q^2)\log\left(\frac{\sqrt{-q^2}\sqrt{4m^2-q^2}+2m^2-q^2}{2m^2}\right)}{\sqrt{-q^2}\sqrt{4m^2-q^2}} + \frac{2M\log\left(\frac{M\sqrt{M^2-4m^2}+2m^2-M^2}{2m^2}\right)}{\sqrt{M^2-4m^2}}}{m^2(q^2 - 4M^2) + M^4}. \tag{115}$$

This result has singularities which are expected from unitarity $M^2 = 4m^2$ and $q^2 = 4m^2$ (check that the expression is regular at $q^2 = 0$).

The result however has also an apparent pole at

$$q^2 = 4m^2 - \frac{(M^2 - 2m^2)^2}{m^2}. \tag{116}$$

Let us check if this is actually a singularity or not. By analyzing the numerator we get a surprising result: it is not a singularity for $M^2 < 2m^2$, but it is a singularity for

$$\text{Anomalous threshold}: \quad s = 4m^2 - \frac{(M^2 - 2m^2)^2}{m^2}, \quad M^2 \geq 2m^2. \tag{117}$$

The original graph indeed has a branch point singularity in this case that starts before $4m^2$! This is an example of a "dragon" that hides in the region of small s, or the so-called *anomalous threshold*. More generally, anomalous thresholds are non-analyticities which do not straightforwardly follow from unitarity. In the context of $2 \to 2$ scattering these are singularities which are of the type $s, t = \text{const}$.

Shall we be very surprised by this result? Thinking back in terms of a form factor, the answer is "not really." Consider for example a positronium, which is a bound state of an electron and a positron. From the S-matrix point of view it is a particle with mass $M = 2m + E_n^{Ps}$, where $E_n^{Ps} < 0$ is the binding energy which can be computed using the non-relativistic Schrödinger equation and it satisfies $E_n^{Ps} \ll m$.

Exercise: Remember from your quantum mechanics course the solution to the quantization of the hydrogen atom problem. Let's say the answer for the energy levels of hydrogen is given and is E_n^H, what are the energy levels for positronium E_n^{Ps}?

Plugging this expression for M into the formula for the anomalous threshold (116) we get that

$$q^2 = 16m|E_n^{Ps}|. \tag{118}$$

Therefore photon sees positronium as an object of size $a_0 = \frac{1}{4\sqrt{m|E_n^{Ps}|}}$. Up to $O(1)$ coefficient this is nothing but the Bohr radius of positronium as expected.

The lesson here is that there is nothing anomalous about anomalous thresholds, but one should be careful in extrapolating unitarity relations away from their original region of validity: new physics maybe hiding there.

An interesting question one might ask is the following: let us start with the triangle graph above with $M = m$ (no anomalous thresholds) and continue it to $M > \sqrt{2}m$ (anomalous threshold present). How does the anomalous threshold come about? The answer is that it hides on the second sheet for $M < \sqrt{2}m$ and then enters to the physical sheet as we continue M past $M = \sqrt{2}m$. This is something general to keep in mind. As we analytically continue s and t away from the values for which we understand analytic properties well we should make sure that no new singularities enter the physical sheet.

Finally, let us consider the total cross section of the positronium-positronium scattering, combining what we just discussed with the previous derivation of the Froissart bound we thus expect that

$$\bar{\sigma}_{\rm tot}^{\rm Ps,Ps\to Ps,Ps}(s) \lesssim (\pi a_0^2) s (\log \frac{s}{s_0})^2, \tag{119}$$

where a_0 is the Bohr radius of the positronium as defined above.

4.6 Landau equations: derivation

Here we briefly review the standard derivation of the Landau equations for the Feynman diagrams. A generic Feynman integral with trivial numerators takes the form

$$F = \int \prod_{j=1}^{L} d^d k_j \int_0^1 \prod_{i=1}^{P} d\alpha_i \frac{\delta(1 - \sum_i \alpha_i)}{\psi^P}, \tag{120}$$

where L are the number of loops, P the number of internal lines and the denominator reads

$$\psi = \sum_{j=1}^{P} \alpha_j (k_j^2 - m_j^2)\,, \tag{121}$$

where the $k_{j>L}$ momenta depend linearly on the loop momenta $k_{j\leq L}$, due to momentum conservation at each vertex.

The integration over the loop momentum can then be readily done and yields

$$F = \int_0^1 \prod_{j=1}^{P} dx_j \, \frac{\delta(1 - \sum_{j=1}^{P} x_j)\, C^{P-2L-2}}{D^{P-2L}}\,, \tag{122}$$

with

$$C = \det a_{ij}\ , \qquad D = \det \begin{pmatrix} a_{ij} & -b_j \\ -b_j & c \end{pmatrix}, \tag{123}$$

where $i, j = 1, \dots, L$ and

$$a_{ij} = \frac{1}{2} \frac{\partial^2 \psi}{\partial k_i \partial k_j}\bigg|_{k=0}\,, \quad b_j = \frac{1}{2} \frac{\partial \psi}{\partial k_j}\bigg|_{k=0}\,, \quad c = \psi|_{k=0}\ . \tag{124}$$

As the integral is analytically continued in the Mandelstam variables, the contour of integration may be deformed smoothly to avoid the singularities. If the contour cannot be deformed by e.g. a *pinch* of two singularities of the integrand then the integral itself becomes singular.

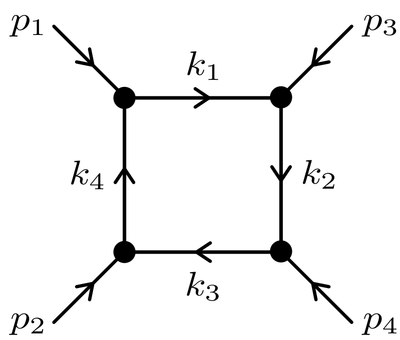

Figure 6: The box graph.

The so-called *leading singularities* occur whenever two (or more) zeros of the denominator coincide.[18] These can be found by solving

$$\frac{\partial\psi}{\partial\alpha_i} = 0 \ , \qquad \frac{\partial\psi}{\partial k_j} = 0 \, . \tag{125}$$

The first condition puts all internal legs on-shell, $k_i^2 = m_i^2$, while the third condition relates momenta belonging to the same loop, l

$$\sum_{i\in l} \alpha_i k_i = 0 \, . \tag{126}$$

An equivalent form of the Landau equations is obtained for representation (122),

$$D = 0 \ , \qquad \frac{\partial D}{\partial\alpha_i} = 0 \, . \tag{127}$$

Note that since $D \propto \alpha_i \frac{\partial D}{\partial\alpha_i}$ is homogeneous, $D = 0$ is automatically satisfied.

There are $P + 2$ variables, s, t and the α parameters, and $P + 1$ Landau equations, which are the P pinch conditions (127) supplemented by the normalization

$$\sum_{i=1}^{P} \alpha_i = 1 \, . \tag{128}$$

These equations may be solved for $\alpha_i(s)$ and $t(s)$, the Landau curve.

4.7 Mandelstam representation: the box graph

Consider an equal mass box integral, see figure 6. In $d = 4$ the result takes the form

$$I(s,t,m^2) = \int_0^1 \prod_{i=1}^{4} d\alpha_i \frac{\delta(1-\sum_{i=1}^4 \alpha_i)}{(s\alpha_1\alpha_3 + t\alpha_2\alpha_4 + m^2(\alpha_1\alpha_2 + \alpha_2\alpha_3 + \alpha_1\alpha_4 + \alpha_3\alpha_4 - 1))^2} \, . \tag{129}$$

The Landau equations take the form

$$\begin{aligned} (s-4m^2)\alpha_3 - 2m^2\alpha_4 &= 0, \\ -2m^2\alpha_3 - (t-4m^2)\alpha_4 &= 0, \end{aligned} \tag{130}$$

[18] There are also end-point singularities, corresponding to pinches at end-points of the integration contour. However, in the context of Feynman integrals, these are also leading singularities of contractions of the original graph [2].

where we used the symmetry of the graph to impose $\alpha_1 = \alpha_3$ and $\alpha_2 = \alpha_4$. To get a nontrivial solution the determinant should be zero which gives the equation for the Landau curve

$$(s-4m^2)(t-4m^2) - 4m^4 = 0. \tag{131}$$

We finally get for the Feynman parameters

$$\alpha_3 = \frac{t-4m^2}{t-2m^2}, \quad \alpha_4 = \frac{m^2}{t-2m^2}. \tag{132}$$

which indeed satisfy $\alpha_3 + \alpha_4 = \frac{1}{2}$. If we send $s \to 4m^2$ or $t \to 4m^2$ we recover the normal threshold solution to the contracted Landau equations, where a pair of Feynman parameters has been set to 0.

It is convenient instead to write the result via *the Mandelstam representation* [20]

$$\begin{aligned} I(s,t,m^2) &= \int_{4m^2}^{\infty} \frac{ds'dt'\rho(s',t')}{(s-s')(t-t')}, \\ \rho(s',t') &= \frac{2\theta((s'-4m^2)(t'-4m^2)-4m^4)}{(s't')^{1/2}\sqrt{(s'-4m^2)(t'-4m^2)-4m^4}}, \end{aligned} \tag{133}$$

where $\theta(x) = 1$ for $x \geq 0$ and 0 otherwise.

Consider a trivial change of variables $s' \to 4m^2 + 2m^2 s'$, $t' \to 4m^2 + 2m^2 t'$ so that we get

$$I(s,t,m^2) = \int_0^{\infty} \frac{ds'dt'}{\sqrt{s'+2}\sqrt{t'+2}} \frac{\theta(s't'-1)}{\sqrt{s't'-1}} \frac{1}{s-4m^2-2m^2s'} \frac{1}{t-4m^2-2m^2t'}. \tag{134}$$

Note that for $0 < s,t < 4m^2$ the integral is manifestly regular. Then we define it by analytic continuation.

At this point it is useful to note that the spectral density kernel admits a very simple Mellin transform

$$\frac{\theta(x-1)}{\sqrt{x-1}} = \int_{-i\infty}^{i\infty} \frac{d\alpha}{2\pi i} \frac{\pi^{1/2}\Gamma(\frac{1}{2}-\alpha)}{\Gamma(1-\alpha)} x^{-\alpha} , \tag{135}$$

where $\text{Re}[\alpha] < \frac{1}{2}$. For $x < 1$ we close the contour to the left and get zero. For $x > 1$ we close the contour to the right and get $\frac{1}{\sqrt{x-1}}$.

Applying this identity to the formula above we note that the integral over s' and t' thus completely factorizes, so that we get

$$\begin{aligned} I(s,t,m^2) &= \int_{-i\infty}^{i\infty} \frac{d\alpha}{2\pi i} \frac{\pi^{1/2}\Gamma(\frac{1}{2}-\alpha)}{\Gamma(1-\alpha)} \tilde{I}(s,m^2,\alpha)\tilde{I}(t,m^2,\alpha), \\ \tilde{I}(s,m^2,\alpha) &\equiv \int_0^{\infty} \frac{ds'}{\sqrt{s'+2}} \frac{(s')^{-\alpha}}{s-4m^2-2m^2s'}. \end{aligned} \tag{136}$$

The integral $\tilde{I}(s,m^2,\alpha)$ can be explicitly computed and it takes the following form

$$\tilde{I}(s,m^2,\alpha) = \frac{i\pi}{\sqrt{2s}m} \frac{(\frac{4m^2-s}{2m^2})^{-\alpha}}{\cos \pi\alpha} + \frac{2^{\frac{1}{2}-\alpha}\Gamma(1-\alpha)\Gamma(\alpha-\frac{1}{2})}{\sqrt{\pi}(s-4m^2)} \, {}_2F_1(1, 1-\alpha, \frac{3}{2}-\alpha, \frac{4m^2}{4m^2-s}). \tag{137}$$

The first term which is purely imaginary for $0 < s < 4m^2$ specifies the branch of ${}_2F_1$ so that the sum is real. This form is particularly useful to analyze the analytic continuation of the integral to $s, t > 4m^2$. Indeed in this case the second term is manifestly real and the continuation of the first term is trivial.

4.7.1 Analytic continuation to $s, t > 4m^2$

There are two inequivalent ways to analytically continue $I(s,t,m^2)$ which we denote $I_{++}(s,t,m^2)$ and $I_{+-}(s,t,m^2)$ (the other two are given by complex conjugation). The sign refers to going above/below the cut in s or t. Correspondingly, we have

$$\tilde{I}_{+}(s,m^2,\alpha) = e^{i\pi\alpha}\frac{i\pi}{\sqrt{2s}m}\frac{(\frac{s-4m^2}{2m^2})^{-\alpha}}{\cos\pi\alpha} + \frac{2^{\frac{1}{2}-\alpha}\Gamma(1-\alpha)\Gamma(\alpha-\frac{1}{2})}{\sqrt{\pi}(s-4m^2)}\ {}_2F_1(1,1-\alpha,\frac{3}{2}-\alpha,\frac{4m^2}{4m^2-s}),$$
$$\tilde{I}_{-}(s,m^2,\alpha) = e^{-i\pi\alpha}\frac{i\pi}{\sqrt{2s}m}\frac{(\frac{s-4m^2}{2m^2})^{-\alpha}}{\cos\pi\alpha} + \frac{2^{\frac{1}{2}-\alpha}\Gamma(1-\alpha)\Gamma(\alpha-\frac{1}{2})}{\sqrt{\pi}(s-4m^2)}\ {}_2F_1(1,1-\alpha,\frac{3}{2}-\alpha,\frac{4m^2}{4m^2-s})\ , \tag{138}$$

where note the appearance of phase in the first term. The second term is manifestly real for $s > 4m^2$.

In this way we get

$$I_{++}(s,t,m^2) = \int_{-i\infty}^{i\infty}\frac{d\alpha}{2\pi i}\frac{\pi^{1/2}\Gamma(\frac{1}{2}-\alpha)}{\Gamma(1-\alpha)}\tilde{I}_{+}(s,m^2,\alpha)\tilde{I}_{+}(t,m^2,\alpha),$$
$$I_{+-}(s,t,m^2) = \int_{-i\infty}^{i\infty}\frac{d\alpha}{2\pi i}\frac{\pi^{1/2}\Gamma(\frac{1}{2}-\alpha)}{\Gamma(1-\alpha)}\tilde{I}_{+}(s,m^2,\alpha)\tilde{I}_{-}(t,m^2,\alpha)\ . \tag{139}$$

Let us ignore for a moment the regular term with ${}_2F_1$ in (138). We immediately see the clear difference between the two

$$I_{++}(s,t,m^2) = -\frac{\pi^2}{\sqrt{st}m^2}\int_{-i\infty}^{i\infty}\frac{d\alpha}{2\pi i}\frac{\pi^{1/2}\Gamma(\frac{1}{2}-\alpha)}{\Gamma(1-\alpha)}\frac{e^{i\ 2\pi\alpha}}{(\cos\pi\alpha)^2}\left(\frac{(s-4m^2)(t-4m^2)}{4m^4}\right)^{-\alpha},$$
$$I_{+-}(s,t,m^2) = -\frac{\pi^2}{\sqrt{st}m^2}\int_{-i\infty}^{i\infty}\frac{d\alpha}{2\pi i}\frac{\pi^{1/2}\Gamma(\frac{1}{2}-\alpha)}{\Gamma(1-\alpha)}\frac{1}{(\cos\pi\alpha)^2}\left(\frac{(s-4m^2)(t-4m^2)}{4m^4}\right)^{-\alpha}\ . \tag{140}$$

The Landau curve is located at $\frac{(s-4m^2)(t-4m^2)}{4m^4} = 1$. The two integrals behave very differently in this limit. Indeed as $\alpha \to -i\infty$ we have

$$\frac{e^{i\ 2\pi\alpha}}{(\cos\pi\alpha)^2} \to 1,$$
$$\frac{1}{(\cos\pi\alpha)^2} \to e^{-2\pi|\alpha|}. \tag{141}$$

This leads to the fact that $I_{+-}(s,t,m^2)$ is regular in the limit $\frac{(s-4m^2)(t-4m^2)}{4m^4} \to 1$ or at the origin of the Landau curve, whereas $I_{++}(s,t,m^2)$ exhibits a square-root singularity $\frac{1}{\sqrt{\frac{(s-4m^2)(t-4m^2)}{4m^4}-1}}$ which is nothing but the spectral density that we started with.

The fact that we got that $I_{+-}(s,t,m^2)$ to be regular is not accidental. If we look at the equation for the Landau curve (131) and think of this as a curve in $\mathbb{C}^2$ we see that the complex part of the Landau surface satisfies

$$\text{Im}s \times \text{Im}t < 0. \tag{142}$$

Therefore if were to find that $I_{+-}(s,t,m^2)$ is singular this singularity would propagate into the complex domain on the physical sheet and would invalidate the Mandelstam representation (133).

4.8 Maximal analyticity

It is clear from the discussion above that unitarity highly constrain the analytic structure of the amplitude. It is tempting to conjecture that analyticity is fully fixed by unitarity on the physical sheet. This conjecture gained some support in the context of scattering of *lightest particles* and the corresponding maximal analyticity is well encapsulated by the Mandelstam representation.

The precise conjecture is the following:

Lighest Particle Maximal Analyticity: The $2 \to 2$ scattering amplitude of the lightest particles in the theory, $T(s,t)$, is analytic on the physical sheet for arbitrary complex s and t, except for potential bound-state poles, a cut along the real axis starting at $s = 4m^2$, and the images of these singularities under the crossing symmetry transformations.

Establishing this hypothesis even within the framework of perturbation theory is an important, open problem in S-matrix theory. Assuming lighest particle maximal analyticity (LPMA), the analytic structure of the $2 \to 2$ amplitude is concisely encapsulated by the Mandelstam representation.[19] From the point of view of our analysis, the nontrivial fact about LPMA is that scattering of lightest particles contains infinitely many subgraphs that by themselves do not respect maximal analyticity. For example, some of the subgraphs that enter the scattering do not admit the Mandelstam representation [20]. For LPMA to hold, embedding these subgraphs inside a larger graph that describes scattering of the lightest particles in the theory should render the complicated singularities of the subgraph harmless on the physical sheet. We have not studied the mechanism of how this happens, and we leave this important question for future work. LPMA is a working assumption in some of the recent explorations of the S-matrix bootstrap, see e.g. [29, 31, 35].

Comment: As a general comment, assuming the Mandelstam representation one can improve various bounds, e.g. (106), and thus try to rule it out experimentally. To the best of our knowledge the experimental data is consistent with the validity of the Mandelstam representation (it would be good if someone checks it carefully!).

5 Dispersion relations

After we introduced some basic notions of S-matrix theory it is time to explore implications of these properties. A very useful tool for that come from studying dispersion relations which nicely encapsulate many of the properties that we have discussed so far.

We start by writing down dispersion relations for scattering of identical particles of mass m. To derive dispersion relations we fix $t < 4m^2$ and consider the integral around ∞

$$\oint_\infty \frac{ds'}{2\pi i} \frac{T(s',t)}{(s-s')(s_1-s')(s_2-s')} = 0, \tag{143}$$

where we used the results of the previous sections, namely (104). In the formula above s_1 and s_2 can depend on t. By closing the contour we pick two type of contributions: residues coming from the poles $\frac{T(s',t)}{(s-s')(s_1-s')(s_2-s')}$ and the discontinuity of the amplitude $T(s,t)$. Via (143), the two

[19] The Mandelstam representation involves an extra assumption that the discontinuity of the amplitude is polynomially bounded on the physical sheet.

contributions should be equal to each other giving

$$\frac{T(s,t)}{(s_1-s)(s_2-s)}+\frac{T(s_1,t)}{(s-s_1)(s_2-s_1)}+\frac{T(s_2,t)}{(s-s_2)(s_1-s_2)}$$
$$+\int_{4m^2}^{\infty}\frac{ds'}{\pi}\frac{T_s(s',t)}{(s-s')(s_1-s')(s_2-s')}+\int_{-\infty}^{-t}\frac{ds'}{\pi}\frac{T_s(s',t)}{(s-s')(s_1-s')(s_2-s')}=0. \tag{144}$$

We can make use crossing by doing the change of variable in the last term $s'\to 4m^2-t-s'$ and recalling that

$$T_s(4m^2-t-s',t)=-T_s(s',t). \tag{145}$$

In this way we get

$$T(s,t)=\frac{1}{(s_2-s_1)}\Big((s-s_2)T(s_1,t)-(s-s_1)T(s_2,t)\Big)$$
$$+(s_1-s)(s_2-s)\int_{4m^2}^{\infty}\frac{ds'}{\pi}T_s(s',t)\left(\frac{1}{(s'-s)(s_1-s')(s_2-s')}+\frac{1}{(s'-u)(u_1-s')(u_2-s')}\right). \tag{146}$$

This formula is known as dispersion relations with two subtractions. It expresses the scattering amplitude as a sum of a linear polynomial in s (the first line) plus the integral over its discontinuity (the second line).

5.1 Bound on chaos

Let us consider subtraction with a double pole $s_1=s_2=2m^2-\frac{t}{2}$

$$\frac{T(s,t)}{(s-2m^2+t/2)^2}=\frac{1}{2\pi i}\oint_{\mathcal{C}_s}\frac{ds'}{s'-s}\frac{T(s',t)}{(s'-2m^2+t/2)^2}=\frac{T(2m^2-t/2,t)}{(s-2m^2+t)^2}-\frac{\partial_s T(2m^2-t/2,t)}{(2m^2-s-t/2)}$$
$$+\frac{1}{\pi}\int_{4m^2}^{\infty}\frac{ds'}{s'-s}\frac{T_s(s',t)}{(s'-2m^2+t/2)^2}+\frac{1}{\pi}\int_{-\infty}^{-t}\frac{ds'}{s'-s}\frac{T_s(s',t)}{(s'-2m^2+t/2)^2}. \tag{147}$$

Note that the point $s=2m^2-\frac{t}{2}$ goes to itself under crossing transformation $s\to 4m^2-s-t$. As a result due to $s-u$ crossing, $\partial_s^{2n+1}T(2m^2-t/2,t)=0$.

After some simple algebraic manipulations, (147) becomes

$$\frac{T(s,t)-T(2m^2-t/2,t)}{(s-2m^2+t/2)^2}=\frac{1}{\pi}\int_{4m^2}^{\infty}ds'\,\frac{T_s(s',t)}{(s'-2m^2+t/2)^2}\left(\frac{1}{s'-s}+\frac{1}{s'-u}\right). \tag{148}$$

This formula is manifestly invariant under crossing $s\to 4m^2-s-t$.

Let us introduce a new variable for complex s, $s=2m^2-t/2+(x+iy)$ we have then

$$\frac{T(s(x,y),t)-T(s(0,0),t)}{(x+iy)^2}=\frac{2}{\pi}\int_{2m^2+t/2}^{\infty}\frac{dx'}{x'}\,\frac{T_s(s(x',0),t)}{x'^2-(x+iy)^2} \tag{149}$$

In particular, we have

$$\frac{T(s(0,0),t)-T(s(0,y),t)}{y^2}=\frac{2}{\pi}\int\limits_{2m^2+t/2}^{\infty}\frac{dx'}{x'}\frac{T_s(s(x',0),t)}{x^2+y^2}=\int\limits_{2m^2+t/2}^{\infty}dx\;\frac{\rho(x,t)}{x^2+y^2}\;, \tag{150}$$

where

$$\rho(x,t)=\frac{2}{\pi x}\,T_s(2m^2-t/2+x,t)\geq 0,\quad 0\leq t\leq 4m^2. \tag{151}$$

Expanding both sides at small y^2 we get the relations between derivatives of the amplitudes and set of moments of a positive density, (with respect to $1/x$). So we can use the Stieltjies moment problem results to write down the set of necessary and sufficient conditions (see Wikipedia).

We can extract from (150) the following local bound on growth of the scattering amplitude. We can set $x=0$ and take a derivative with respect to y to get rid of the subtraction term. In this way we get

$$\partial_y T(s(y),t)=-2y\int\limits_{2m^2+t/2}^{\infty}dx\;\rho(x,t)\frac{x^2}{(x^2+y^2)^2}\leq 0 \tag{152}$$

and hence

$$\text{Bound on chaos}:\quad -4\leq y\partial_y\log\partial_y T(s(y),t)-1=-4\frac{\int\limits_{2m^2+t/2}^{\infty}dx\;\frac{\rho(x,t)x^2}{(x^2+y^2)^2}\frac{y^2}{x^2+y^2}}{\int\limits_{2m^2+t/2}^{\infty}dx\;\frac{\rho(x,t)x^2}{(x^2+y^2)^2}}\leq 0. \tag{153}$$

This condition says that locally

$$\text{if}\quad T(s(y),t)\sim y^{\alpha}\quad\text{then}\quad -2\leq\alpha\leq 2. \tag{154}$$

An example of a theory that saturates the upper bound is famously given by a weakly coupled gravitational theory. In this case the graviton exchange in the t-channel gives $T(s,t)\sim G_N\frac{s^2}{t}$. Of course here we talk about gapped theories but we can still ask about the variation of the Weinberg-Witten question: is it possible to have emergent gravity within the framework of QFT? In the context of the present discussion we can ask if it is possible to saturate the bound for $\alpha=2$.

Alternatively from (149) we see that (here we can use that $T(s(0,0),t)$ is real)

$$\begin{aligned}(xy)\times\mathrm{Im}\,[T(s(x,y),t)]=&\;\frac{2}{\pi}\int\limits_{2m^2+t/2}^{\infty}\frac{dx'}{x'}\;T_s(s(x',0),t)\,\mathrm{Im}\frac{xy(x+iy)^2}{x'^2-(x+iy)^2} \\ =&\;\frac{4}{\pi}\int\limits_{2m^2+t/2}^{\infty}dx'\;\frac{T_s(s(x',0),t)\,(xy)^2x'}{[x'^2-(x+iy)^2][x'^2-(x-iy)^2]}\geq 0\end{aligned} \tag{155}$$

Again we can bound the local behavior of the imaginary part. We get

$$4\leq y\partial_y\log\mathrm{Im}\,[T(s(x,y),t)]-1=\frac{\int\limits_{2m^2+t/2}^{\infty}dx'\;T_s(s(x',0),t)\;\frac{-4x'y^2(x^2+x'^2+y^2)}{[x'^2-(x+iy)^2]^2[x'^2-(x-iy)^2]^2}}{\int\limits_{2m^2+t/2}^{\infty}dx'\;T_s(s(x',0),t)\;\frac{x'}{[x'^2-(x+iy)^2][x'^2-(x-iy)^2]}}\leq 0\;. \tag{156}$$

Which again states that the imaginary part does not grow faster than linearly.

5.2 Dispersive couplings and null constraints

Let us next do something outrageous, namely set $m = 0$ and (if we are very brave) include dynamical gravity. Remarkably, even though many of the arguments (especially everything based on QFT axioms) go out of the window we believe that dispersion relations still hold. Let us also for simplicity assume that the number of subtractions stays 2.[20]

A fruitful way to think about the dispersion relations of the type (143) is

$$\text{IR} + \text{UV} = 0. \tag{157}$$

In this formula the IR part corresponds to the part of the dispersive contour which can be computed using the low-energy effective theory. The UV part expresses everything that we do not know about the theory (however general principles such as unitarity still apply!).

For example we can ask about the structure of the higher-derivative operators in the theory. In the absence of high energy experimental data this sounds like a very reasonable strategy to make theoretical progress. Instead of talking about higher derivative operators we can introduce the low-energy expansion of the amplitude

$$T(s,t) = -g^2\left(\frac{1}{s}+\frac{1}{t}+\frac{1}{u}\right) - g_0 + g_2(s^2+t^2+u^2) + g_3 stu + \sum_{p,q=0,2p+3q\geq 4}^{\infty} g_{2p+3q,q}\sigma_2^p\sigma_3^q\,, \tag{158}$$

where we introduced

$$\sigma_k \equiv (-1)^k \frac{s^k+t^k+u^k}{k}. \tag{159}$$

Note that $\sigma_3 = -stu$.

Here g_i are called Wilson coefficients and they can be related to the higher derivative operators in the Lagrangian. In writing (158) we have neglected coming from loops of massless particles, treating this requires care and this is something that has not been done very carefully yet. An important comment about (158) is that it is manifestly fully crossing symmetric.

Consider next dispersion relations (143) with $s_1 = 0$ and $s_2 = -t$. The result takes the form

$$-\frac{T(s,t)}{s(s+t)} + \text{Res}_{s'=0,-t}\frac{T(s',t)}{(s-s')s'(t+s')} - \int_{m_{gap}^2}^{\infty}\frac{ds'}{\pi}\frac{T_s(s',t)}{s'(s'+t)}\left(\frac{1}{s-s'}+\frac{1}{u-s'}\right) = 0. \tag{160}$$

We can rewrite this as follows

$$T(s,t) = -g^2\left(\frac{1}{s}+\frac{1}{t}+\frac{1}{u}\right) + f_0(t) + sf_1(t) + s(s+t)\int_{m_{gap}^2}^{\infty}\frac{ds'}{\pi}\frac{T_s(s',t)}{s'(s'+t)}\left(\frac{1}{s'-s}+\frac{1}{s'-u}\right) \tag{161}$$

Let us next compare dispersion relation (161) to the low-energy expansion (158). We see that it is natural to split the set of all couplings $g_{i,j}$ into two classes: *dispersive* and *non-dispersive*. Dispersive couplings are those that come from the integral over discontinuity (these are forces generated by the UV). Non-dispersive couplings are those that do not admit such a representation. In the example above g^2 and g_0 are non-dispersive couplings. All other couplings are dispersive.

[20]For gravitational theories this should hold for $d > 7$ only, for $5 \leq d \leq 7$ the discussion requires some inessential modification. For $d = 4$ the question is open.

The simplest set of constraints one derives from dispersion relation takes the following form [37]

$$\mu_n = \frac{1}{2}\frac{1}{n!}\partial_s^{2n}\left(T(s,t) + g^2\left(\frac{1}{s}+\frac{1}{t}+\frac{1}{u}\right)\right)|_{s,t=0} = \int_{m_{gap}^2}^{\infty}\frac{ds'}{\pi}\frac{T_s(s',0)}{(s')^{2n-1}} \geq 0, \tag{162}$$

where we used unitarity which states that $T_s(s',0) > 0$. We therefore see that in the forward limit we naturally get a one-dimensional moment problem with very well-known consequences, see e.g. [38].

The basic development of the last year is a simple observation that going away from $t = 0$ leads to a two-dimensional moment problem [39]. To see this, let us note that the imaginary part admits a decomposition in terms of partial waves with non-negative coefficients

$$T_s(s',t) = \sum_J n_J^{(d)} \text{Im} f_J(s) P_J^{(d)}(1 + \frac{2t}{s'}), \quad \text{Im} f_J(s') \geq 0. \tag{163}$$

We can think of $\text{Im} f_J(s) \sim \sum_X \lambda_{\phi\phi X^2}$ where X are all possible states of spin J and mass $m^2 = s$. Expanding (163) around $t = 0$ leads to the following series

$$P_J^{(d)}(1+x) = 1 + \frac{\mathcal{J}^2 x}{d-2} + \frac{\mathcal{J}^2(2-d+\mathcal{J}^2)}{2d(d-2)}x^2 + ..., \tag{164}$$

where we introduced $\mathcal{J}^2 = J(J+d-3)$ for the d-dimensional quadratic Casimir of the rotation group $SO(d)$. Therefore by studying expansion in both s and t we study two-dimensional moments of the type

$$\mu_{m,n} = \sum_J \int \frac{ds'}{(s')^m}\rho_J(s')\mathcal{J}^{2n}, \tag{165}$$

where $\rho_J(s') = n_J^{(d)} \text{Im} f_J(s') > 0$. This two-dimensional moment problem that arises in the study of the dispersion relations was dubbed *the EFThedron* [42]. It leads to much more complicated set of bounds on couplings that follow from (162).

There is an interesting twist to this story! Observe the sharp contrast between the low-energy expansion (158) and the dispersive representation of the amplitude (160) is that in the latter form only the $s \leftrightarrow u$ crossing is manifest. In other words, crossing symmetry $T(s,t) = T(t,s)$ leads to the following sum rule

$$\text{Dispersive crossing}: \quad s\Big(\text{Res}_{s'=0,-t}\frac{T(s',t)}{(s-s')s'(t+s')} + \int_{m_{gap}^2}^{\infty}\frac{ds'}{\pi}\frac{T_s(s',t)}{s'(s'+t)}\left(\frac{1}{s-s'}+\frac{1}{u-s'}\right)\Big)$$
$$= t\Big(\text{Res}_{s'=0,-s}\frac{T(s',s)}{(t-s')s'(s+s')} + \int_{m_{gap}^2}^{\infty}\frac{ds'}{\pi}\frac{T_s(s',s)}{s'(s'+s)}\left(\frac{1}{t-s'}+\frac{1}{u-s'}\right)\Big). \tag{166}$$

While the expression (166) does represent a nontrivial crossing equation it is still not very neat because it involves the terms $\text{Res}_{s'=0,-t}\frac{T(s',t)}{(s-s')s'(t+s')}$ which acquire contributions from infinitely many Wilson coefficients in (158).

At least in the context of *the tree-level EFTs* there is a nice way to get rid of this little nuisance. To do this let us go back to the dispersive representation of the amplitude (161) and consider its expansion around $s, t = 0$. To get rid of the subtractions we start with terms proportional to s^2.

Consider for example the following coupling in the expansion of $T(s,t) \sim g_{4,0}(\frac{s^2+t^2+u^2}{2})^2$. We can write this coupling in two ways

$$12 g_{4,0} = \partial_s^2 \partial_t^2 \hat{T}(s,t)|_{s,t=0} = \partial_s^3 \partial_t \hat{T}(s,t)|_{s,t=0}, \tag{167}$$

where we have introduced

$$\hat{T}(s,t) \equiv T(s,t) + g^2 \left(\frac{1}{s} + \frac{1}{t} + \frac{1}{u} \right). \tag{168}$$

We therefore get the following constraint on the discontinuity of the amplitude

$$\left(\partial_s^2 \partial_t^2 - \partial_s^3 \partial_t\right) \int_{m_{gap}^2}^{\infty} \frac{ds'}{\pi} \frac{T_s(s',t)}{s'(s'+t)} \left(\frac{1}{s-s'} + \frac{1}{u-s'} \right) \Big|_{s,t=0} = 0. \tag{169}$$

This is the simplest example of what became known as *null constraints.* Null constraints by definition are the consequences of dispersive crossing applied to tree-level EFTs. In the presence of logarithms dispersive crossing still definitely makes sense but extracting its consequences is more complicated. For example, the simplest clean dispersive crossing would be applied to $\partial_s^2 \partial_t^2 \hat{T}(s,t)$ which would be free of subtraction terms.

The basic idea behind null constraints is that they relate low spin contributions to high spin contributions and they have been used to derive two-sided bounds on the *dispersive couplings*, see e.g. [40] and [41] for the detailed derivation of this claim. Differently, we can think of dispersive crossing as taking a particular slice of the EFThedron.

Exercise: Which couplings are dispersive and which are not in the Standard Model coupled to gravity? Are any of the recent developments phenomenologically interesting? To answer this question we need to understand how the story above generalizes when particles with spin are present in the theory.

5.3 Superconvergence

The phenomenon of superconvergence is the following fact:

Spin makes couplings dispersive

To make this precise let us consider scattering of gravitons. For simplicity we restrict kinematics to four-dimensions (to make the polarization analysis simple and to psychologically connect to reality).

Gravitons are labeled by helicity and the simplest (elastic or non-helicity flipping amplitude) is given by the following expression

$$\begin{aligned} \mathcal{M}_4(1^+,2^-,3^-,4^+) &= (\langle 23\rangle[14])^4 f(s,u), \\ f_{\rm GR}(s,u) &= \left(\frac{\kappa}{2}\right)^2 \frac{1}{stu} + \dots . \end{aligned} \tag{170}$$

More generally this function satisfies crossing

$$f(s,u) = f(u,s), \tag{171}$$

and let's assume that the amplitude satisfies the same bound

$$\lim_{|t|\to\infty} \mathcal{M}_4(1^+,2^-,3^-,4^+) < |t|^2, \quad s<0. \tag{172}$$

The crucial fact is that $(\langle 23\rangle[14])^4 \sim |t|^4$, therefore we have $t^2 f(s,-s-t) \to 0$. This leads to extra sum rules of the type

$$\oint_\infty dt t^n f(s,-s-t) = 0, \quad n = 0,1,2. \tag{173}$$

These sum rules mean that all terms in the low-energy expansion of $f(s,u)$ are dispersive (including the Einstein term)! This is something special about scattering of gravitons: their attraction comes fully from integrating out the UV modes.

As an example let us write down the representation analogous to the one in the previous section (see [43] for details)

$$\begin{aligned} f(t,-s-t) = \oint \frac{ds'}{2\pi i}\frac{f(t,-s'-t)}{s-s'} &= \left(\frac{\kappa}{2}\right)^2 \frac{1}{stu} + |\beta_{R^3}|^2 \frac{tu}{s} - |\beta_\phi|^2 \frac{1}{s} \\ &- \int_{m_{\rm gap}^2}^\infty \frac{dm^2}{\pi}\Bigg(\sum_{J=0}^\infty \frac{1+(-1)^J}{2}\frac{\rho_J^{++}(m^2) d_{0,0}^J(1+\frac{2t}{m^2})}{m^8}\frac{1}{s-m^2} \\ &\qquad + \sum_{J=4}^\infty \frac{\rho_J^{+-}(m^2) d_{4,4}^J\left(1+\frac{2t}{m^2}\right)}{(t+m^2)^4}\frac{1}{-s-t-m^2}\Bigg). \end{aligned} \tag{174}$$

Note that the formula above is not manifestly crossing-symmetric $f(t,u) = f(u,t)$, therefore we have the constraints of dispersive crossing which becomes null constraints if we neglect the loops.

Exercise: Work out the implications of the dispersive relations for the possible UV completions of the Standard Model coupled to gravity in flat space.

6 Bootstraps

In this section we discuss various incarnations of the S-matrix bootstrap: old and new. These will go beyond general bounds based on dispersion relations that we discussed in the previous section. We will explicitly specify what are the extra assumptions in each case. We also comment on the connections to the CFT bootstrap.

6.1 The Froissart-Gribov formula

Extra assumption: Extended region of analyticity.

The Froissart-Gribov formula is a representation of the partial wave coefficients in terms of the discontinuity of the amplitude. It has multiple applications and, in particular, it allows us to analytically continue partial wave coefficients in spin.

Let us introduce the Gegenbauer Q-functions. These are given by the second linearly independent solution of the second order Casimir equation (62). They are uniquely fixed by their asymptotic behavior

$$\lim_{|z|\to\infty} Q_J^{(d)}(z) = \frac{c_J^{(d)}}{z^{J+d-3}} + \dots \,, \tag{175}$$

where $c_J^{(d)}$ is a normalization constant. The corresponding Q-function is

$$Q_J^{(d)}(z) = \frac{c_J^{(d)}}{z^{J+d-3}} {}_2F_1\left(\frac{J+d-3}{2}, \frac{J+d-2}{2}, J+\frac{d-1}{2}, \frac{1}{z^2}\right) \,. \tag{176}$$

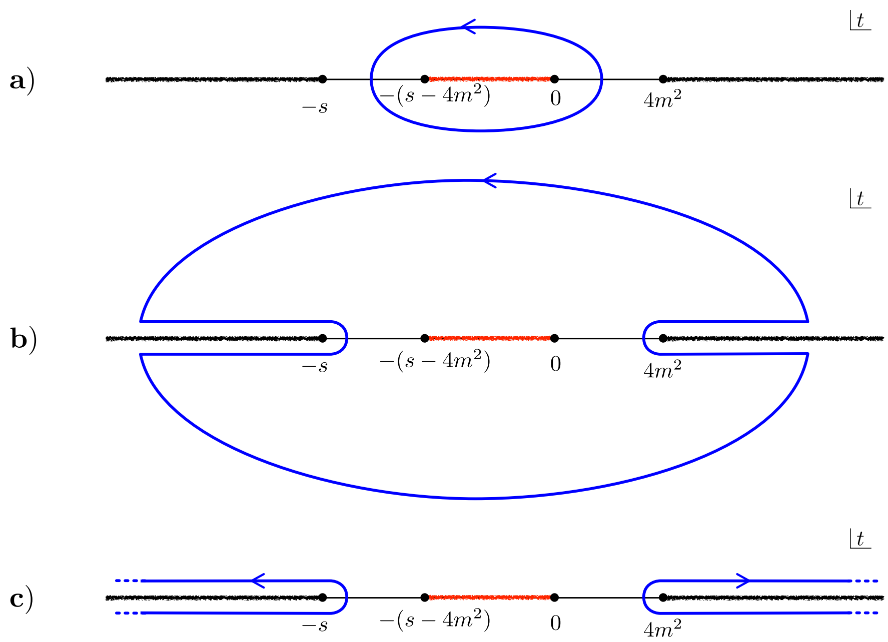

Figure 7: **a**. The partial wave projection integral (180) is a contour integral (in blue) that circles around the cut of the Q-function, between $t = 0$ and $t = -(s - 4m^2)$, (in red). **b**. We partially open up the contour. Sometimes this representation for partial waves is called the truncated Froissart-Gribov formula. The advantage of this representation is that we only use a finite amount of extended analyticity that has not been rigorously proven. **c**. We open the contour all the way to infinity and arrive at the usual Froissart-Gribov formula (181) with two integrations of the discontinuity of the amplitude along the t-channel and u-channel cuts (in black).

Our convention is

$$c_J^{(d)} = \frac{\sqrt{\pi}\Gamma(J+1)\Gamma(\frac{d-2}{2})}{2^{J+1}\Gamma(J+\frac{d-1}{2})} \ . \tag{177}$$

The Q-function has a cut running between $z = -1$ and $z = 1$. The fact that there are only two independent solutions to the Casimir equation means that the discontinuity of Q can be expressed in terms of Q and P. The precise relation takes the form

$$\text{Disc}_z(z^2-1)^{\frac{d-4}{2}} Q_J^{(d)}(z) = -\frac{\pi}{2}(1-z^2)^{\frac{d-4}{2}} P_J^{(d)}(z) \ , \qquad z \in [-1,1] \ , \tag{178}$$

or equivalently (for integer J)

$$Q_J^{(d)}(z) = \frac{1}{2}\int_{-1}^{1} dz' \left(\frac{1-z'^2}{z^2-1}\right)^{\frac{d-4}{2}} \frac{P_J^{(d)}(z')}{z-z'} \ . \tag{179}$$

We can then plug (B.4) into the partial wave coefficient (66) as

$$f_J(s) = \mathcal{N}_d \oint_{[-1,1]} \frac{dz}{2\pi i} \left(z^2-1\right)^{\frac{d-4}{2}} Q_J^{(d)}(z) T(s,t(z)) \ , \tag{180}$$

where the integral is counterclockwise around the interval $z \in [-1,1]$. By blowing up the contour, we get two integrals along the t- and the u-channel cuts, see figure 7

$$f_J(s) = \frac{\mathcal{N}_d}{\pi}\left[\int_{z_1}^{\infty} dz(z^2-1)^{\frac{d-4}{2}} Q_J^{(d)}(z) T_t(s,t(z)) + \int_{-\infty}^{-z_1} dz(z^2-1)^{\frac{d-4}{2}} Q_J^{(d)}(z) T_u(s,u(z))\right] , \tag{181}$$

where

$$z_1 \equiv z|_{t=4m^2} = 1 + \frac{8m^2}{s-4m^2} , \tag{182}$$

and we have assumed that $s > 4m^2$, so that the t channel cut runs from $z_1 = z_1 > 1$ to infinity. Here we have dropped the contributions of the arcs at infinity. This is justified for large enough spin $J > J_0(s)$ using (B.1), where $J_0(s)$ is the Regge intercept

$$\lim_{|t|\to\infty} |T(s,t)| < |t|^{J_0(s)} . \tag{183}$$

We can now use crossing to simplify (181). We change the integration variable for the u-channel integral from z to $-z$. Crossing symmetry implies that $T_u(s,u(z)) = -T_t(s,t(-z))$, where we have used that $z(u) = -z$. Under this change of variables

$$(z^2-1)^{\frac{d-4}{2}} \to (-1)^{d-4}(z^2-1)^{\frac{d-4}{2}} , \qquad Q_J^{(d)}(z) \to Q_J^{(d)}(-z) = (-1)^{J+3-d} Q_J^{(d)}(z) . \tag{184}$$

We get that $f_J = 0$ for odd J. For even J we get

$$\boxed{f_J(s) = \frac{2\mathcal{N}_d}{\pi}\int_{z_1}^{\infty} dz\,(z^2-1)^{\frac{d-4}{2}} Q_J^{(d)}(z) T_t(s,t(z)) , \qquad \text{Re}J > J_0(s)} . \tag{185}$$

As opposed to (66), the Froissart-Gribov representation of the partial waves (185) is suitable for analytic continuation in J. It follows from the Carlson's theorem that this analytic continuation is the unique continuation that does not grow too fast at large J. The Froissart-Gribov integral (185) converges as long as Re$J > J_0(s)$ thanks to (183) and (B.1).

This integral is written for $s > 4m^2$. As s approaches the threshold from above $s - 4m^2 \to 0^+$, the lower end of the integral is pushed to infinity, $z_1 \to \infty$. To analyze $f_J(s)$ in this limit, it is useful to use (B.1) and to switch back to an integral over t. In that way one finds

$$f_J(s) = \frac{2\mathcal{N}_d}{\pi} c_J^{(d)} \left(\frac{s-4m^2}{2}\right)^J \int_{4m^2}^{\infty} \frac{dt}{t^{J+1}} T_t(4m^2,t)\left(1 + O\left((s-4m^2)/t\right)\right) . \tag{186}$$

This integral should be understood as follows. The large t contribution is finite because $|\text{Im}_t T(4m^4,t)| < |T(4m^2,t)| < t^{J_0(4m^2)}$ and $J > J_0(4m^2)$ by assumption. If the integrand diverges at some finite t, and in particular as $t - 4m^2 \to 0^+$, then we should step back and write it as a contour integral of $T(4m^2,t)$ around the cut, which is manifestly finite.

Comments:

- The coefficients that multiply $(s-4m^2)^J$ and measures in the low-energy experiments are known as *scattering lengths*. From the formula above we again encounter a one-dimensional moment problem since $T_t(4m^2,t) \geq 0$. For more on this see [10].

- In deriving the FG formula we had to assume extra analyticity beyond what has been proven. In CFTs the analogous formulas (the so-called *Lorentzian inversion formula* [45]) does not require extra assumptions.

- *Nuclear democracy* is a dynamical assumption that analyticity in spin continues to $J = 0$, namley that the Froissart-Gribov formula maybe we some simple subtractions should correctly reproduce the spectrum of $J = 0$ particles: *All particle poles are Regge poles.* There has been an interesting discussion of something similar happening in CFTs recently [46].

6.1.1 Neat subtractions

There is a convenient way to write down dispersion relations without arbitrary subtraction constants. Consider the following split of the amplitude in spin (recall that sum goes over even spins only)

$$T(s,t) = \sum_{J=0}^{J_0-2} n_J^{(d)} f_J(s) P_J^{(d)}(z) + \sum_{J=J_0}^{\infty} n_J^{(d)} f_J(s) P_J^{(d)}(z). \tag{187}$$

Let us now plug the Froissart-Gribov formula (185) for the high spin tail

$$\sum_{J=J_0}^{\infty} n_J^{(d)} f_J(s) P_J^{(d)}(z) = \sum_{J=J_0}^{\infty} n_J^{(d)} P_J^{(d)}(z) \frac{2\,\mathcal{N}_d}{\pi} \int_{z_1}^{\infty} d\hat{z}\, (\hat{z}^2-1)^{\frac{d-4}{2}} Q_J^{(d)}(\hat{z}) T_t(s,t(\hat{z})). \tag{188}$$

We can use the identity[21]

$$\frac{1}{2}\left(\frac{1}{\hat{z}-z}+\frac{1}{\hat{z}+z}\right) = \mathcal{N}_d(\hat{z}^2-1)^{\frac{d-4}{2}} \sum_{J=0, J-even}^{\infty} n_J^{(d)} P_J^{(d)}(z) Q_J^{(d)}(\hat{z}) \tag{189}$$

to rewrite the sum as follows

$$\mathcal{N}_d(\hat{z}^2-1)^{\frac{d-4}{2}} \sum_{J=J_0}^{\infty} n_J^{(d)} P_J^{(d)}(z) Q_J^{(d)}(\hat{z}) = \frac{1}{2}\left(\frac{1}{\hat{z}-z}+\frac{1}{\hat{z}+z}\right) - \mathcal{N}_d(\hat{z}^2-1)^{\frac{d-4}{2}} \sum_{J=0}^{J_0-2} n_J^{(d)} P_J^{(d)}(z) Q_J^{(d)}(\hat{z}). \tag{190}$$

In this way we get the following representation for the amplitude

$$T(s,t) = \sum_{J=0}^{J_0-2} n_J^{(d)} f_J(s) P_J^{(d)}(z) \tag{191}$$

$$- \frac{2}{\pi} \int_{z_1}^{\infty} d\hat{z} \left(\mathcal{N}_d\, (\hat{z}^2-1)^{\frac{d-4}{2}} \sum_{J=0}^{J_0-2} n_J^{(d)} P_J^{(d)}(z) Q_J^{(d)}(\hat{z}) + \frac{\hat{z}}{z^2-\hat{z}^2} \right) T_t(s,t(\hat{z}))\ . \tag{192}$$

As expected the expression in the brackets behaves as $\frac{1}{\hat{z}^{J_0+1}}$ at large $\hat{z}$. Namely, the first $1/\hat{z}^{2n+1}$ terms with $n < J_0/2$ cancels between the two terms in the bracket.

---

[21] Note that this is nothing but $\frac{1}{2}$ times the dispersion integral for the sum of the the t and the u channels. Namely, $\frac{1}{2}\left(\frac{1}{\hat{z}-z}+\frac{1}{\hat{z}+z}\right) = \frac{1}{2}\left(\frac{1}{\hat{z}-z(t)}+\frac{1}{\hat{z}-z(u)}\right)$.

6.2 The Aks theorem and bound on inelasticity

Extra assumptions: Extended analyticity.

Let us now ask the following question: is it possible to have scattering without production when the number of spacetime dimensions is $d > 2$? In $d = 2$ this is definitely possible and the corresponding theories are integrable [47]. In $d > 2$ the answer to this question is negative: given scattering, production is necessary.

We now review the elegant argument that establishes this for scalar particles by Aks [21]. It states that scattering implies particle production in $d > 2$. Namely, provided that $T_{2\to 2} \neq 0$, also $T_{2\to n} \neq 0$ with $n > 2$. The theorem applies to any crossing symmetric scalar scattering amplitude in $d \geq 3$ that satisfies extended analyticity in a finite region above the leading Landau curve.

To derive the result of Aks, let us therefore assume that we have a nontrivial scattering amplitude $T(s,t)$, but $T_{2\to n}$ are identically zero for $n > 2$. This implies that elastic unitarity holds for any $s > 4m^2$. Let us write it down explicitly

$$f_J(s+i0) - f_J(s-i0) = i\frac{(s-4m^2)^{\frac{d-3}{2}}}{\sqrt{s}} f_J(s+i0) f_J(s-i0), \tag{193}$$

where in our previous discussion $f_J(s) \equiv f_J(s+i0)$. From the results of the previous section we see that (193) actually holds for any complex J with $\mathrm{Re}[J] > J_0(s)$!

There is a nice way to express elastic unitarity at complex J which is known as *the Mandelstam equation*

$$\rho(s,t) = \frac{(s-4m^2)^{\frac{d-3}{2}}}{4\pi^2 (4\pi)^{d-2}\sqrt{s}} \int_{z_1}^{\infty} d\eta' \int_{z_1}^{\infty} d\eta''\, T_t(s+i0,\eta') T_t(s-i0,\eta'')\, \mathrm{Disc}_z K_d(z,\eta',\eta'') , \tag{194}$$

where recall that $\rho(s,t)$ is the double discontinuity of the amplitude. Here, the lower limit of integration is the point where the t channel cut starts (182). The discontinuity of the kernel, for $\eta'\eta'' > 0$ and $z > 1$ is given by

$$\mathrm{Disc}_z K_3(z,\eta',\eta'') = 4\pi^2 \delta(z-\eta_+) \frac{\sqrt{z^2-1}}{\eta_+ - \eta_-} ,$$
$$\mathrm{Disc}_z K_{d\geq 4}(z,\eta',\eta'') = \frac{4\pi^{\frac{d+1}{2}}}{\Gamma(\frac{d-3}{2})} \Theta(z-\eta_+) \frac{(z^2-1)^{\frac{4-d}{2}}}{(z-\eta_-)^{\frac{5-d}{2}} (z-\eta_+)^{\frac{5-d}{2}}} \geq 0 . \tag{195}$$

What is the relation between (194) and (193)? In fact it is very easy

$$(193): \quad \int_{z_1}^{\infty} dz\, (z^2-1)^{\frac{d-4}{2}} Q_J(z) \quad (194). \tag{196}$$

This relies on some very special property of the Mandelstam kernel (195), see [35] for details.

Finally, what is the advantage of (194) compared to (193)? It is crossing! Indeed, we have

$$\rho(s,t) = \rho(t,s). \tag{197}$$

Moreover, using the Mandelstam equation and partial wave expansion, one can show that whenever elastic unitarity applies we have

$$\text{Nontrivial scattering}: \quad \rho(s,t) > 0 , \quad \frac{16m^2 s}{s-4m^2} < t \leq 4m^2 \frac{(3s+4m^2)^2}{(s-4m^2)^2}. \tag{198}$$

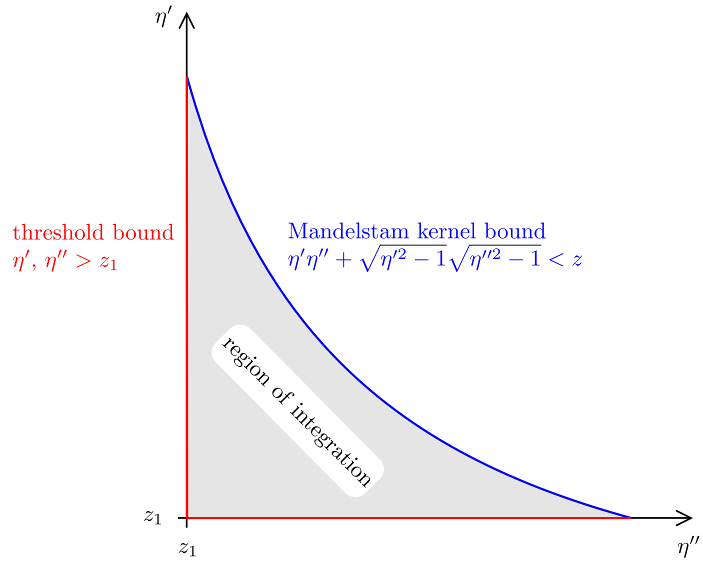

Figure 8: The region of integration in equation (194). As s or t approaches the Landau curve from above, the integration region shrinks to zero. As a result, the double spectral density vanishes below the Landau curve $z = 2z_1^2 - 1$.

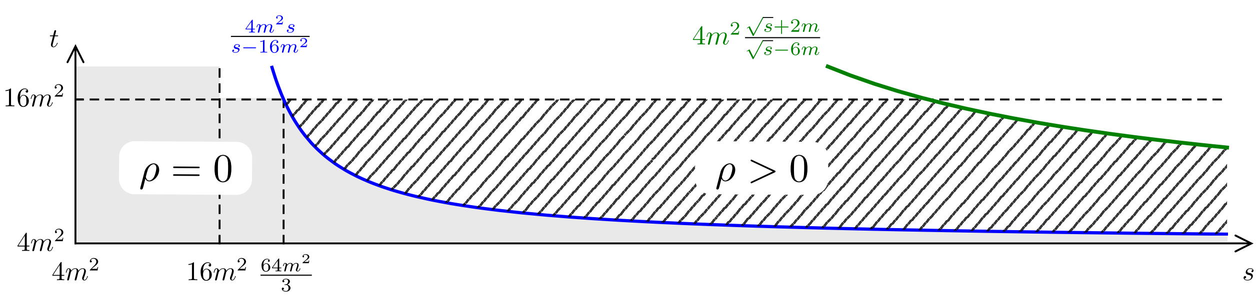

Figure 9: The regime of positivity of the double spectral density $\rho(s,t)$ in the elastic strip $4m^2 < t < 16m^2$, (dashed region). This is the region above the first Karplus curve $s_1(t)$ (in blue) and below the curve $t = 4m^2\frac{\sqrt{s}+2m}{\sqrt{s}-6m}$ (in green). There is an identical positive region in the crossed strip of $4m^2 < s < 16m^2$, see (198).

This result follows from unitarity and applying partial wave expansion to (194).

Now comes the beautiful argument of Aks: nontrivial scattering (or (198)) is not consistent with zero production (which is validity of (194) beyond $s < 16m^2$) due to crossing (197) (which implies vanishing of $\rho(s,t)$ in the region (198)).

Note that the argument above also implies that we must have four-particle production. That is because the crossed region of positivity starts at $s_1(16m^2) = \frac{64}{3}m^2 < 36m^2$, it is enough to assume that $T_{2\to4} = 0$ to reach a contradiction. One can wonder if having $T_{2\to4}$ is enough to fix the problem, or $T_{2\to2n}$ with $n > 2$ are also necessary? To address this question, we can then proceed via crossing.[22] By unitarity of the $4\to4$ amplitude, $\text{Im}T_{4\to4} \sim |T_{2\to4}|^2$, non-vanishing $T_{2\to4}$ implies that we have a non-vanishing $T_{4\to4}$ amplitude. Applying crossing this becomes $T_{2\to6}$. Continuing this recursion we conclude that all $T_{2\to2n}$ should be non-zero. Therefore, not only scattering implies production but it requires all possible production (here we assumed Z_2 symmetry so that only an

[22]Note that crossing has been only rigorously proven within the standard QFT framework for $2\to2$ (and $2\to3$) amplitudes [12].

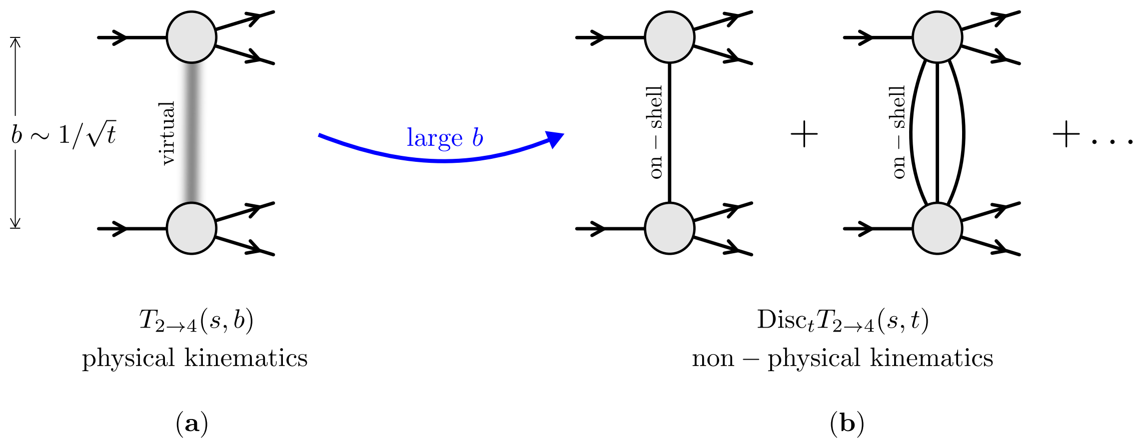

Figure 10: **a**. We consider a scattering experiment at fixed impact parameters b. This process is controlled by the exchanged momentum $t \sim \frac{1}{b^2} > 0$. **b**. At large impact parameters the amplitude can be organized as a sum over on-shell particles exchange. The dominant contribution comes from the exchanged of the lightest on-shell particle.

even number of particles is present in the final state).

Gribov's theorem: Interestingly, elastic unitarity continued in spin constrains the high-energy behavior of the amplitude. Indeed, consider $t \to \infty$ limit and imagine that the amplitude (and its discontinuity) behave as $T \sim t^{J_0(s)}$. Via the Froissart-Gribov formula this leads to $f_J(s) \sim \frac{1}{J - J_0(s)}$. Gribov's theorem is the statement that $J_0(s)$ cannot be real for $4m^2 < s < s_{MP}$ (whenever elastic unitarity (193) holds). The reason is trivial: as we take $J \to J_0(s)^+$ the RHS of (193) has a double pole, whereas the LHS at most single pole. This is of course a contradiction. The simplest resolution is that $J_0(s) \in \mathbb{C}$.

6.3 Bounding Inelasticity

There is another, more intuitive way to think about the result of Aks and necessity of particle production in higher dimensions. Ideally, one would like to take a discontinuity of the $2 \to 4$ amplitude that is given by a product of two $2 \to 2$ amplitudes. For physical kinematics however, such a discontinuity only exist for the $3 \to 3$ setup. Instead, let us discuss impact parameter scattering, which is only possible in $d \geq 3$.

As reviewed for example in appendix E of [34], the effect of going to the impact parameter space is the same as continuing the conjugate momentum invariant to the unphysical kinematics. It follows that inelasticity in a gapped theory cannot be exactly zero at very large impact parameters. To see this, decompose the four particles in the final state into a pair of dipoles. Then consider a scattering in which the two dipoles in the final state, as well as the pair of incoming particles in the initial state, are separated by a finite distance b in the transverse space, see figure 10.a. Unitarity becomes, in the impact parameter space, the expansion in Yukawa potential suppressed terms $T_{2\to4}(s,b) \sim T^2_{2\to2} \times e^{-bm} +$ [multi-particle $\sim e^{-nbm}$]. At large separation, this expansion is dominated by the one-particle exchange e^{-bm}, while the multi-particle corrections are further exponentially suppressed.[23] In that way, a non-trivial (analytically continued) four-point amplitude imply a non-trivial $2 \to 4$ amplitude. We would then like to bound the $2 \to 4$ amplitude from below.

[23]In terms of partial waves, large impact parameter scattering corresponds to the large spin limit and therefore we expect to have inelasticity at large spin which will be analyzed in detail in the sections below.

There is a convenient way of bounding the integrated discontinuity of $T_{2\to 4}$ in the kinematical regime of figure 10.b from below. Let us start with the the square of $T_{2\to 4}$ that appears in the discontinuity of the $2 \to 2$ amplitude, $T_s(s,t)$

$$T_s^{\text{inel},2\to 4}(s,t) \equiv \frac{1}{2}\frac{1}{4!}\int \prod_{i=1}^{4} \frac{d^{d-1}\vec{q}_i}{(2\pi)^{d-1}(2E_{\vec{q}_i})}(2\pi)^d \delta^d(p_1+p_2-\sum_{i=1}^{4} q_i) T_{2\to 4}^{(+)}(p_1,p_2|q_i) T_{2\to 4}^{(-)}(q_i|p_3,p_4) \ . \tag{199}$$

By construction, the unitarity integral in the right-hand side of (199) depends only on s and t. For physical scattering we consider $s > 16m^2$ and $t < 0$.

We would like next to analytically continue (199) to the unphysical Martin-Mahoux region discussed above

$$4m^2 < t < 16m^2 \ , \qquad \frac{16m^2 t}{t-4m^2} < s < 4m^2 \frac{(3t+4m^2)^2}{(t-4m^2)^2} \ . \tag{200}$$

We also would like to consider $16m^2 < s < 36m^2$ to focus on the $T_{2\to 4}$ amplitude. This condition together with (200) imply $\frac{36m^2}{5} < t$.

After taking a discontinuity in t and using crossing symmetry of the double spectral density, we arrive at the following schematic form[24]

$$\begin{aligned}\rho(s,t) &= \frac{(t-4m^2)^{\frac{d-3}{2}}}{4\pi^2(4\pi)^{d-2}\sqrt{t}} \int_{\bar{z}_1}^{\infty} d\eta' \int_{\bar{z}_1}^{\infty} d\eta'' \, \text{Disc}_s T_{2\to 2}^{(+)}(t,\eta') \, \text{Disc}_s T_{2\to 2}^{(-)}(t,\eta'') \times \text{Disc}_{\bar{z}} K_d(\bar{z},\eta',\eta'') \\ &= \int d\text{LIPS}_4 \times \, \text{Disc}_t T_{2\to 4}^{(+)} \, \text{Disc}_t T_{2\to 4}^{(-)} \times K_{\text{Mandelstam}}^{2\to 4} \ ,\end{aligned} \tag{201}$$

where in the formula above we switched to $\bar{z} = 1 + \frac{2s}{t-4m^2}$ and $\bar{z}_1 = 1 + \frac{8m^2}{t-4m^2}$. Here in the right-hand side each $\text{Disc}T_{2\to 4}$ contains a delta-function that puts the exchanged particle in figure 10 on-shell. The phase space integral $d\text{LIPS}_4$ should be understood in terms of the analytic continuation a-la Mandelstam.

6.4 The Dragt bootstrap: large J is simple

Extra assumptions: Extended analyticity.
Result: Large J expansion of partial waves is mapped to the low-energy physics.

There is an analytic S-matrix bootstrap that maps near-threshold physics in one channel to the large spin physics in the other channel. It is in some sense a simple generalization of what is depicted in figure 10 and goes as follows [22].

Let us write down a few key formulas:

- Low-energy/threshold expansion for $\sigma_s = \frac{s}{4m^2} - 1 \ll 1$. Solution to elastic unitarity takes the form

$$\frac{1}{f_J(s+i\epsilon)} = b_J(\sigma_s) - \frac{i}{2}\frac{(4m^2\sigma_s)^{\frac{d-3}{2}}}{\sqrt{s}} \times \begin{cases} 1 & d \text{ even} \\ \frac{i}{\pi}[\log \sigma_s - i\pi] & d \text{ odd} \end{cases} \ , \tag{202}$$

  where $b_J(\sigma_s)$ is a real analytic and single-valued function in some finite neighborhood around the origin, except for potentially isolated singularity at $\sigma_s = 0$.

[24] Note that $T_{2\to 2}$ is not present in the second line of (201) since the region $4m^2 < t < 16m^2$, $16m^2 < s < 36m^2$ lies below the leading t-channel Landau curve $t = 16m^2 s/(s-4m^2)$.

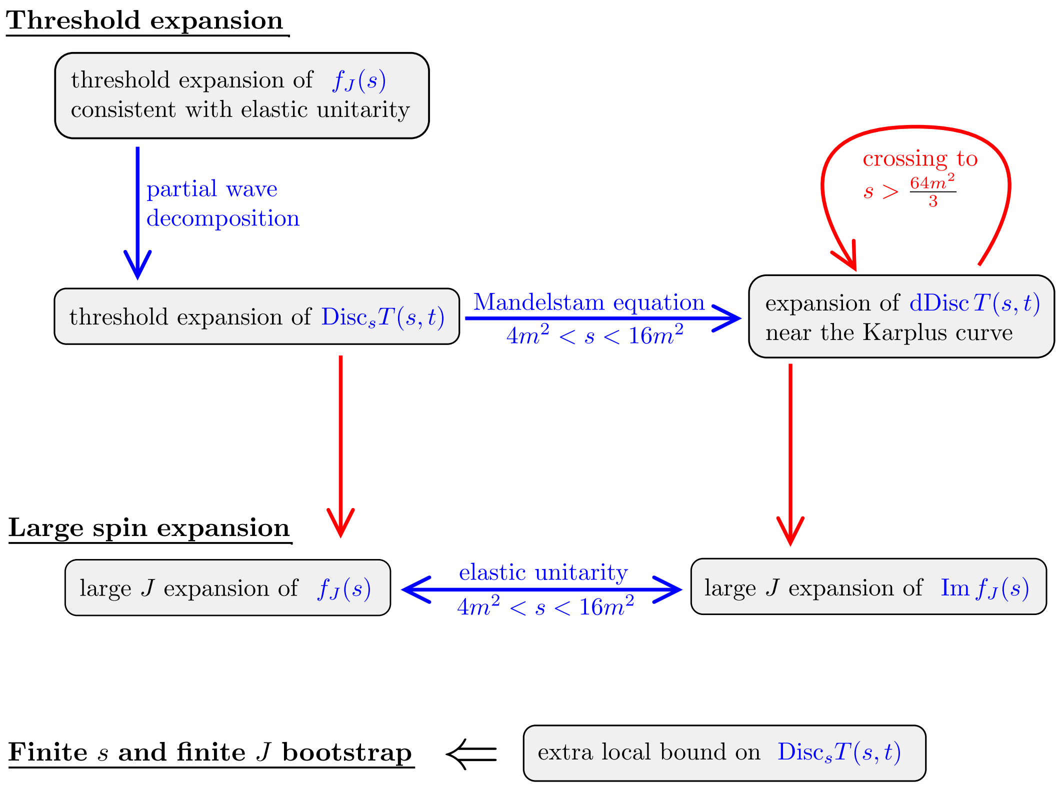

Figure 11: The analytic S-matrix bootstrap (or the Dragt boostrap).

- Via the Mandelstam equation this leads to the threshold expansion for $\rho(s,t)$ close to the leading Landau curves (where we also used crossing).
- The Froissart-Gribov formula then turns the near-threshold expansions into the large J expansion of partial waves. This is based on the following simple fact

$$Q_J^{(d)}(z) = 2^{d-4}\sqrt{\pi}\, \frac{\Gamma\left(\frac{d-2}{2}\right)}{J^{\frac{d-3}{2}}} \frac{\lambda(z)^{-J}}{(\lambda(z)^2-1)^{\frac{d-3}{2}}} \left(1 + \mathcal{O}(1/J)\right) \ , \tag{203}$$

The weakness of this approach is lack of good control over the corrections at finite s and finite J.

Comment: A very similar scheme works for CFTs, indeed in this case instead of the threshold expansion it is the lightcone expansion that gets mapped to the large spin J expansion [48, 49]. Until recently it was also not clear how to control this scheme at finite J. Recently however such a proposal has been made in [50]. Presumably, a similar improvement should work here (but this has not been done).

6.5 The fixed point method

Extra assumptions: Mandelstam representation. Multi-particle physics is given.

One possible answer to the question "what are the nonperturbative scattering amplitudes?" is somewhat trivially:

$$\boxed{\text{Nonperturbative scattering amplitudes are fixed points of the unitarity equations.}} \tag{204}$$

Of course unitarity equations (44) include infinitely many scattering amplitudes $T_{m\to n}$, however they admit natural simplifications in theories of massive particles. Indeed, as we discussed before if energy is such that no particle production is possible, say for $4m^2 \leq s < 16m^2$ unitarity leads to a non-linear equation on the $2 \to 2$ amplitude. Similarly, taking $16m^2 \leq s < 36m^2$ only $T_{2\to 2}$ and $T_{2\to 4}$ enter the unitarity relations. At this point to close the system of equations we need to include more $T_{m\to n}$ into the considerations, or alternatively we can keep talking about $T_{2\to 2}$ only and introduce some parameterization for our ignorance of the multi-particle contributions. Until today all attempts to do nonperturbative S-matrix bootstrap followed the latter path: including multi-particle amplitudes into considerations looks like a complicated (if not unsurmountable) task.

Let us describe this idea in the simplest possible higher-dimensional setting ($d > 2$). We assume that the scattering amplitude admits Mandelstam representation without subtractions

$$T(s,t) = \frac{1}{\pi^2} \int_{4m^2}^{\infty} \frac{ds' dt' \ \rho(s',t')}{(s'-s)(t'-t)} + \frac{1}{\pi^2} \int_{4m^2}^{\infty} \frac{du' dt' \ \rho(u',t')}{(u'-u)(t'-t)} + \frac{1}{\pi^2} \int_{4m^2}^{\infty} \frac{ds' du' \ \rho(s',u')}{(s'-s)(u'-u)}, \tag{205}$$

where $\rho(s,t)$ is the double spectral density.

As we briefly discussed the double spectral density acquires its support along the Landau curves which (at least for the case of the lightest mass scattering) form an hierarchical structure. For example a recent study [27] reveals the following structure of the leading Landau curves.

It is therefore very natural to write down the following split

$$\rho(s,t) = \rho_{el}(s,t) + \rho_{el}(t,s) + \rho_{MP}(s,t), \tag{206}$$

where $\rho_{MP}(s,t)$ stands for multi-particle contribution, and $\rho_{el}(s,t)$ is the elastic double spectral density that satisfies the Mandelstam equation.

In the Atkinson program, $\rho_{MP}(s,t)$ is an input and does not change in the process of iterations. It encodes the contribution of the multi-particle scattering which thus does not participate in the bootstrap scheme. Elastic unitarity is imposed via the Mandelstam equation

$$\rho_{el}(s,t) = \frac{(s-4m^2)^{\frac{d-3}{2}}}{4\pi^2 (4\pi)^{d-2} \sqrt{s}} \int_{z_1}^{\infty} d\eta' \int_{z_1}^{\infty} d\eta'' T_t^{(+)}(s,\eta') T_t^{(-)}(s,\eta'') \, \mathrm{Disc}_z K_d(z,\eta',\eta'') \ , \tag{207}$$

where $T_t^{(\pm)}$ is the t-channel discontinuity of the amplitude.

The iteration process now goes as follows:

1. Start with a fixed $\rho_{MP}(s,t)$ which has only support above the leading multi-particle Landau curve.
2. Compute the discontinuity $T_t(s,t)$ of the amplitude using (205).
3. Compute $\rho_{el}(s,t)$ using (194).
4. Update $\rho(s,t)$ using (206) and repeat steps 2-4.
5. Compute the full amplitude $T(s,t)$ and check inelastic unitarity.

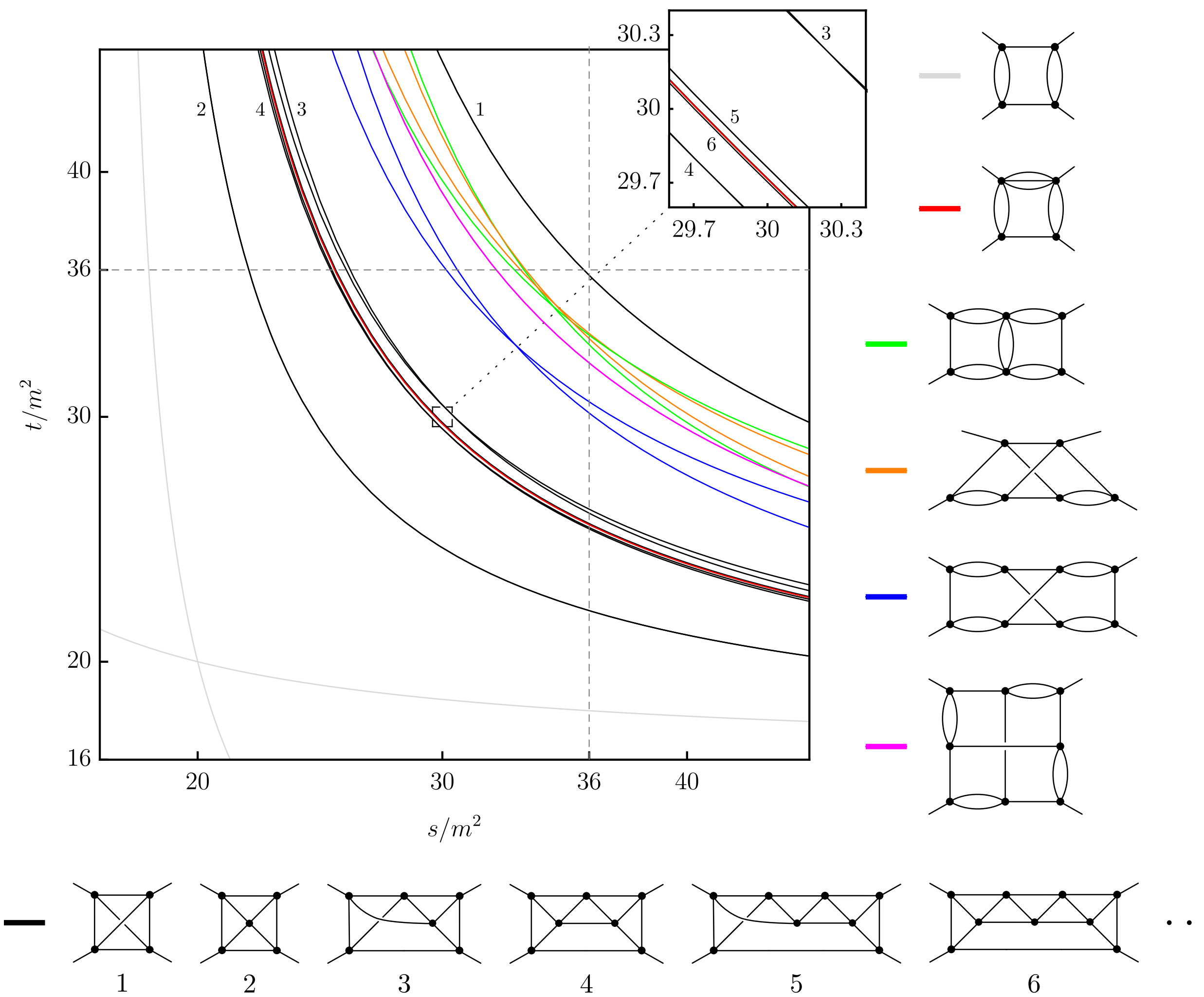

Figure 12: The leading Landau curves for the lightest particle scattering.

In this scheme unitarity serves as a mapping that generates the amplitude at the next iteration step. Atkinson has proven in [24] that if $\rho_{MP}(s,t)$ satisfies some additional constraints, this mapping is contracting and therefore it converges and the solution is unique (see REF for a pedagogical introduction [51]). This allows one to construct a function $T(s,t)$ that satisfies properties (elastic+inelastic) unitarity, crossing, and analyticity. Until today this is the only proposed algorithm to impose these conditions.

6.5.1 The strip model approximation

In the old days people tried something called the strip model approximation. It states that:

$$\boxed{\text{Strip model approximation: } T(s,t) \text{ is small unless } s,\, t, \text{ or } u \text{ is close to } 4m^2.} \tag{208}$$

This is an ad hoc proposal of dealing with the multi-particle physics in the scheme above. The basic hope is that ρ_{MP} should be somehow not very relevant and therefore one starts with the amplitude ansatz that incorporates a few Regge trajectories (say observed in the experiment) and then runs the iteration scheme above, where the ad hoc modification for the Mandelstam equation is also proposed (to enlarge it to the multi-particle region).

The name for the strip model comes from the fact that effectively all the action and the support of the double spectral density is effectively restricted to the elastic strips $4m^2 < s,t,u < 16m^2$. It would be useful to have a critical assessment of this scheme from the standpoint of our current understanding of both QCD and the S-matrix bootstrap. To the best of my knowledge this has not been done.

In fact, it seems that in QCD the opposite is true and the amplitude is very big for $s,t \gg 4m^2$ [32]!

6.6 Primal bootstrap

Extra assumptions: Maximal analyticity. Multi-particle physics is arbitrary. Elastic unitarity and the multi-particle threshold structure is not respected.

Here we briefly review the setup of [29]. The basic idea underlying this approach can be stated as follows: given that we cannot properly control the multi-particle physics, we should focus on the questions for which it is maximally irrelevant. A natural candidate is the low-energy coupling that can be defined as the value of the amplitude at the crossing-symmetric point $s = t = u = \frac{4}{3}m^2$. Further explorations indeed provided evidence that this question is not very sensitive to the multi-particle structure of the amplitude.

One then writes an ansatz for the expansion of the scattering amplitude which is linear in unknown real parameters α_{abc}

$$T(s,t) = \sum_{a,b,c=0} \alpha_{abc}\rho_s^a \rho_t^b \rho_u^c + \text{extra}\Big|_{u=4m^2-s-t} , \tag{209}$$

where, the function $\rho_s \equiv \frac{\sqrt{4m^2-s_0}-\sqrt{4m^2-s}}{\sqrt{4m^2-s_0}+\sqrt{4m^2-s}}$ maps the complex s-plane minus the s-channel cut to the unit circle and the point s_0 to the origin. Here, the extra terms may be added to make some particular properties of the amplitude manifest. Their presence or absence depends on the particular problem at hand. Crossing symmetry is imposed by demanding that the coefficients α_{abc} are permutation-invariant. Finally, the relation $s+t+u=4m^2$ leads to a redundancy in the basis of coefficients that can be addressed systematically.

To approximate an amplitude using the ansatz (209), the sum is truncated such that

$$a+b+c \leq N_{\max} \ . \tag{210}$$

Given a finite $N_{\max}$, unitary in the form

$$|S_J(s)| \leq 1 \ , \qquad s \geq 4m^2 \ , \qquad J \in 2\mathbb{Z}_+ \ , \tag{211}$$

is imposed over a finite grid of points and for spins that are truncated by some maximal value $J \leq J_{\max}(N_{\max})$. As shown in [29], remarkably unitarity in the form of (211) can be restated as a semidefiniteness condition as follows. We write for physical J and s

$$S_J(s) = 1 + i\vec{\alpha}\cdot\vec{f}_J(s) \ , \tag{212}$$

where $\hat{f}_J(s)$ are kinematical objects and all the dynamical information is in the coefficients $\vec{\alpha}$. The condition (211) can be then rewritten as a semi-definitedness condition for the matrix

$$M \equiv \begin{pmatrix} 1+\vec{\alpha}\cdot\mathrm{Re}\vec{f}_J(s) & 1-\vec{\alpha}\cdot\mathrm{Im}\vec{f}_J(s) \\ 1-\vec{\alpha}\cdot\mathrm{Im}\vec{f}_J(s) & 1-\vec{\alpha}\cdot\mathrm{Re}\vec{f}_J(s) \end{pmatrix} \succcurlyeq 0 \ . \tag{213}$$

At this point one can maximize numerically some quantity linear in the α-parameters by imposing unitarity in the form (213) over the chosen grip in s and for $J \leq J_{\max}(N_{\max})$. For example, in [29] the "coupling", $T(\frac{4m^2}{3}, \frac{4m^2}{3})$, is maximized. If a certain maximization task reliably saturates as a function of $N_{\max}$ we stop the process and trivially extrapolate to $N_{\max}=\infty$ to get the actual bound on the space of physical S-matrices.

6.7 Dual bootstrap

Extra assumptions: None!

While in all other approaches extra assumptions that go beyond what has been proven are necessary there is a way to derive bounds without extra assumptions. In fact, it has been known how to derive such rigorous bounds since 70's! Recently this approach has been revisited in the paper by Andrea Guerrieri and Amit Sever [30]. I refer the reader to this paper (and references therein) and to the lectures by Andrea Guerrieri at the Bootstrap School 2021 for further details.

A Derivation of the Unitarity Kernel

The kernel $\mathcal{P}_d(z,z',z'')$ Recall that z' and z'' are cosine of the angles between $\vec{n}$ in (52) and the vectors $\vec{p}_1$ and $\vec{p}_3$, (51). They are related to the coordinates in the Sudakov decomposition of the unit vector $\vec{n}$

$$\vec{n} = \alpha\frac{\vec{p}_1}{|\vec{p}_1|} + \beta\frac{\vec{p}_3}{|\vec{p}_3|} + \vec{n}_\perp \ , \qquad \vec{n}_\perp\cdot\vec{p}_1 = \vec{n}_\perp\cdot\vec{p}_3 = 0 \tag{A.1}$$

as

$$z' = \alpha + \beta z \ , \qquad z'' = \beta + \alpha z \ , \qquad \alpha = \frac{z' - zz''}{1-z^2} \ , \qquad \beta = \frac{z'' - zz'}{1-z^2} \ . \tag{A.2}$$

In term of these coordinates, the angular integration in (52) reads

$$\begin{aligned}\int d^{d-2}\Omega_{\vec{n}} &= 2\int d^{d-1}\vec{n}\,\delta(\vec{n}^2-1)\\ &= 2\sqrt{1-z^2}\int d\alpha\, d\beta\, d^{d-3}\vec{n}_\perp \delta(\vec{n}_\perp^2 + \alpha^2+\beta^2+2\alpha\beta z - 1)\\ &= 2\sqrt{1-z^2}\int d\alpha\, d\beta\, \frac{\Theta(1-\alpha^2-\beta^2-2\alpha\beta z)}{(1-\alpha^2-\beta^2-2\alpha\beta z)^{\frac{5-d}{2}}}\int d^{d-3}\vec{n}_\perp\delta(\vec{n}_\perp^2-1),\\ &= \sqrt{1-z^2}\mathrm{Vol}_{S^{d-4}}\int d\alpha\, d\beta\, \frac{\Theta(1-\alpha^2-\beta^2-2\alpha\beta z)}{(1-\alpha^2-\beta^2-2\alpha\beta z)^{\frac{5-d}{2}}} \ ,\end{aligned} \tag{A.3}$$

where $\mathrm{Vol}_{S^{d-4}} = \frac{2\pi^{(d-3)/2}}{\Gamma(\frac{d-3}{2})}$. The above formula is only true in $d \geq 4$. In $d=3$ we have

$$\int d\Omega_{\vec{n}} = 2\int d^2\vec{n}\,\delta(\vec{n}^2-1) = 2\sqrt{1-z^2}\int d\alpha\, d\beta\, \delta(\alpha^2+\beta^2+2\alpha\beta z - 1) \ , \tag{A.4}$$

which can be also obtained as a distributional limit from (A.3) when $d \to 3$. By plugging the relation (A.2) into (A.3) and (A.4), we arrive at (53).

B Q-functions

Let us introduce the Gegenbauer Q-functions. These are given by the second linearly independent solution of the second order Casimir equation (62). They are uniquely fixed by their asymptotic behavior

$$\lim_{|z|\to\infty} Q_J^{(d)}(z) = \frac{c_J^{(d)}}{z^{J+d-3}} + \dots \ , \tag{B.1}$$

where $c_J^{(d)}$ is a normalization constant. The corresponding Q-function is

$$Q_J^{(d)}(z) = \frac{c_J^{(d)}}{z^{J+d-3}}\,{}_2F_1\left(\frac{J+d-3}{2}, \frac{J+d-2}{2}, J+\frac{d-1}{2}, \frac{1}{z^2}\right) \ . \tag{B.2}$$

Our convention is

$$c_J^{(d)} = \frac{\sqrt{\pi}\Gamma(J+1)\Gamma(\frac{d-2}{2})}{2^{J+1}\Gamma(J+\frac{d-1}{2})} \ . \tag{B.3}$$

The Q-function has a cut running between $z=-1$ and $z=1$. The fact that there are only two independent solutions to the Casimir equation means that the discontinuity of Q can be expressed in terms of Q and P. The precise relation takes the form

$$\mathrm{Disc}_z (z^2-1)^{\frac{d-4}{2}} Q_J^{(d)}(z) = -\frac{\pi}{2}(1-z^2)^{\frac{d-4}{2}} P_J^{(d)}(z) \ , \qquad z\in[-1,1] \ , \tag{B.4}$$

or equivalently (for integer J)

$$Q_J^{(d)}(z) = \frac{1}{2}\int_{-1}^{1} dz' \left(\frac{1-z'^2}{z^2-1}\right)^{\frac{d-4}{2}} \frac{P_J^{(d)}(z')}{z-z'} \ . \tag{B.5}$$

C Martin's extension of the Lehmann ellipse

Recall that the Lehmann ellipse shrinks with energy $t_L(s) \sim \frac{C}{s}$ as $s \to \infty$. The basic point of Martin's extension is that such a shrinking of the analyticity domain in t is not consistent with unitarity and in fact we have analyticity for $|t| < R$ where R does not depend on s.

Here we sketch the basic argument postponing further technical details to the original papers [3], [6] . We consider elastic scattering $A, B \to A, B$ and the basic new ingredient compared to the analysis done before Martin is *unitarity*.

More precisely, the assumptions of the original paper [3] are:

1. existence of dispersion relations for $-t_M \leq t \leq 0$ where the amplitude in s is analytic outside of the unitarity cuts;

2. for physical s the amplitude and its absorptive part are analytic inside the Lehmann ellipse;

3. from [13] and [14] it follows that in the neighbourhood of any point s_0 (outside the unitarity cuts) and $-t_M \leq t_0 \leq 0$ there is analyticity in both s and t in

$$|s - s_0| < \eta(s_0, t_0), \quad |t - t_0| < \eta(s_0, t_0). \tag{C.1}$$

A priori $\eta(s_0, t_0)$ can go to zero as $s_0 \to \infty$. Our task is to show that this does not happen.

Let us for simplicity ignore the left cut and subtractions. Therefore we assume that the amplitude is analytic in the cut s-plane with the cut starting from $s \geq (m_A + m_B)^2$. Unitarity

$$\text{Disc}_s T(s, z) = \sum_J n_J^{(d)} \text{Im} f_J(s) P_J^{(d)}(z) \ , \quad \text{Im} f_J(s) \geq 0, \tag{C.2}$$

where $z = 1 + \frac{2t}{s - 4m^2}$ implies that

$$\partial_z^n \text{Disc}_s T(s, z)\Big|_{z=1} \geq 0, \tag{C.3}$$

$$\Big|\partial_z^n \text{Disc}_s T(s, z)\Big|_{-1 \leq z \leq 1} \leq \partial_z^n \text{Disc}_s T(s, z)\Big|_{z=1}. \tag{C.4}$$

In the formulas above we made use of the basic properties of the Legendre polynomials and the fact that $\text{Im} f_J(s) \geq 0$.

Step 1

Consider an unphysical real $s_1 < (m_A + m_B)^2$ (below the cut). Using property 3 the amplitude is analytic in some neighbourhood $|t| < R$. We have at this point for the amplitude

$$\Big|(\frac{d}{dt})^n T(s_1, 0)\Big| \leq \frac{M n!}{R^n}. \tag{C.5}$$

Here we assumed that $T(s_1, Re^{i\phi})$ is finite otherwise we take $R - \epsilon$.

Step 2

Let's write down dispersion relations (unsubtracted, without the left cut)

$$T(s, t) = \frac{1}{\pi} \int_{(m_A + m_B)^2}^{\infty} \frac{ds'}{\pi} \frac{\text{Disc}_s T(s', t)}{s' - s}. \tag{C.6}$$

Martin next argues that we can differentiate under the dispersion relations so that

$$(\frac{d}{dt})^n T(s_1,0) = \frac{1}{\pi}\int_{(m_A+m_B)^2}^{\infty} \frac{ds'}{\pi} \frac{(\frac{d}{dt})^n \text{Disc}_s T(s',0)}{s'-s_1}. \tag{C.7}$$

Let us discuss the first derivative at $t=0$

$$\frac{d}{dt}T(s_1,0) = \lim_{\tau\to 0,\tau>0} \frac{T(s_1,0)-T(s_1,-\tau)}{\tau} = \frac{1}{\pi}\lim_{\tau\to 0,\tau>0}\int_{(m_A+m_B)^2}^{\infty} \frac{ds'}{\pi}\frac{[\text{Disc}_s T(s',0)-\text{Disc}_s T(s',-\tau)]/\tau}{s'-s_1}. \tag{C.8}$$

We split the integral over s' and we use (C.3) (together with a trivial positivity of $\frac{1}{s'-s_1}$) to get

$$\frac{d}{dt}T(s_1,0) \geq \frac{1}{\pi}\lim_{\tau\to 0,\tau>0}\int_{(m_A+m_B)^2}^{x} \frac{ds'}{\pi}\frac{[\text{Disc}_s T(s',0)-\text{Disc}_s T(s',-\tau)]/\tau}{s'-s_1} \tag{C.9}$$

Using analyticity inside the large Lehmann ellipse we can differentiate under the integral and we get

$$\frac{d}{dt}T(s_1,0) \geq \frac{1}{\pi}\int_{(m_A+m_B)^2}^{x} \frac{ds'}{\pi}\frac{\frac{d}{dt}\text{Disc}_s T(s',0)}{s'-s_1}. \tag{C.10}$$

This holds for arbitrary x and recall that $\frac{d}{dt}\text{Disc}_s T(s',t) \geq 0$. Therefore the limit $x\to\infty$ exists and we get

$$\frac{d}{dt}T(s_1,0) \geq \frac{1}{\pi}\int_{(m_A+m_B)^2}^{\infty} \frac{ds'}{\pi}\frac{\frac{d}{dt}\text{Disc}_s T(s',0)}{s'-s_1}. \tag{C.11}$$

Let us now get the opposite estimate. We again split the integral into $\int_{(m_A+m_B)^2}^{x}$ and $\int_x^\infty$

$$\int_x^\infty \frac{ds'}{\pi}\frac{[\text{Disc}_s T(s',0)-\text{Disc}_s T(s',-\tau)]/\tau}{s'-s_1} = \int_x^\infty \frac{ds'}{\pi}\frac{[\frac{d}{dt}\text{Disc}_s T(s',-\tau(s'))}{s'-s_1} \leq \int_x^\infty \frac{ds'}{\pi}\frac{[\frac{d}{dt}\text{Disc}_s T(s',0)}{s'-s_1}, \tag{C.12}$$

where in the first step we used Rolle's theorem? $(0\leq\tau(s')\leq\tau)$ and in the second step unitarity (C.3). To sum up we conclude that

$$\frac{d}{dt}T(s_1,0) = \frac{1}{\pi}\int_{(m_A+m_B)^2}^{\infty} \frac{ds'}{\pi}\frac{\frac{d}{dt}\text{Disc}_s T(s',0)}{s'-s_1}. \tag{C.13}$$

We then combine (C.7) with (C.5) to get

$$|\frac{1}{\pi}\int_{(m_A+m_B)^2}^{\infty} \frac{ds'}{\pi}\frac{(\frac{d}{dt})^n\text{Disc}_s T(s',0)}{s'-s_1}| \leq \frac{Mn!}{R^n}. \tag{C.14}$$

Step 3

We next continue (C.7) in s_1 away from $s_1 < (m_A+m_B)^2$. We then use

$$\left|\frac{(\frac{d}{dt})^n\text{Disc}_s T(s',0)}{s'-s}\right| \leq B(s,s_1)\frac{(\frac{d}{dt})^n\text{Disc}_s T(s',0)}{s'-s_1}, \tag{C.15}$$

where

$$B(s, s_1) \equiv \max_{s' \geq (m_A+m_B)^2} \Big| \frac{s'-s_1}{s'-s} \Big|. \tag{C.16}$$

In this way we conclude that

$$\Big| (\frac{d}{dt})^n T(s,0) \Big| \leq B(s,s_1) \frac{M n!}{R^n}, \tag{C.17}$$

which shows that series

$$T(s,t) = \sum_n \frac{1}{n!} (\frac{d}{dt})^n T(s,0) \tag{C.18}$$

converges for $|t| < R$ where R does not depend on s. Note that $B(s+i\epsilon, s_1) \sim \frac{s}{\epsilon}$.

Further comments:

- Here we have not explained the origin of property 3 which led to the crucial inequality (C.5). At first glance we have not found the discussion of this in [14], but [13] does precisely that using dispersion relations and analyticity inside the Lehmann ellipse.
- We have ignored the left cut and subtractions.
- We have not found what R is for a given scattering process.
- We have not explained how to formulate this discussion using the BES analyticity domain (for which we cannot go to the harmless region $s < (m_A + m_B)^2$).

Martin claims to address all these issues in [6] at least for the pion-pion scattering (and some of them in [3]).

References

[1] J. T. Cushing, "Theory construction and selection in modern physics: The S matrix," Cambridge, UK: Univ. Pr. (1990) 409 p

[2] R. J. Eden, P. V. Landshoff, D. I. Olive and J. C. Polkinghorne, "The analytic S-matrix,"

[3] A. Martin, "Extension of the axiomatic analyticity domain of scattering amplitudes by unitarity. 1.," Nuovo Cim. A **42** (1965), 930-953 doi:10.1007/BF02720568

[4] R. J. Eden, "Theorems on high energy collisions of elementary particles," Rev. Mod. Phys. **43** (1971), 15-35 doi:10.1103/RevModPhys.43.15

[5] G. Sommer, "Present state of rigorous analytic properties of scattering amplitudes," Fortsch. Phys. **18** (1970) 577. doi:10.1002/prop.19700181102

[6] A. Martin, "Scattering Theory: Unitarity, Analyticity and Crossing," Lect. Notes Phys. **3** (1969), 1-117 doi:10.1007/BFb0101043

[7] E. T. Whittaker, G. N. Watson, "A course of modern analysis," Cambridge university press, 1996.

[8] M. Froissart, "Asymptotic behavior and subtractions in the Mandelstam representation," Phys. Rev. **123** (1961), 1053-1057 doi:10.1103/PhysRev.123.1053

[9] F. J. Yndurain, "Absolute bound on cross-sections at all energies and without unknown constants," Phys. Lett. B **31** (1970), 368-370 doi:10.1016/0370-2693(70)90198-X

[10] F. J. Yndurain, "Rigorous constraints, bounds, and relations for scattering amplitudes," Rev. Mod. Phys. **44** (1972) 645. doi:10.1103/RevModPhys.44.645

[11] Y. S. Jin and A. Martin, "Number of Subtractions in Fixed-Transfer Dispersion Relations," Phys. Rev. **135** (1964), B1375-B1377 doi:10.1103/PhysRev.135.B1375

[12] J. Bros, H. Epstein and V. Glaser, "A proof of the crossing property for two-particle amplitudes in general quantum field theory," Commun. Math. Phys. **1** (1965) no.3, 240-264 doi:10.1007/BF01646307

[13] H. Lehmann, "Analytic properties of scattering amplitudes in two variables in general quantum field theory," Commun. Math. Phys. **2** (1966) no.1, 375-383 doi:10.1007/bf01773361

[14] J. Bros, H. Epstein and V. J. Glaser, "Some rigorous analyticity properties of the four-point function in momentum space," Nuovo Cim. **31** (1964), 1265-1302 doi:10.1007/BF02733596

[15] S. Mandelstam, "Determination of the pion - nucleon scattering amplitude from dispersion relations and unitarity. General theory," Phys. Rev. **112** (1958), 1344-1360 doi:10.1103/PhysRev.112.1344

[16] D. Iagolnitzer, "Scattering in quantum field theories: The Axiomatic and constructive approaches," Princeton, USA: Univ. Pr. (1993) 290 p

[17] S. Weinberg, "The Quantum theory of fields. Vol. 1: Foundations,"

[18] D. Iagolnitzer and H. P. Stapp, "Macroscopic causality and physical region analyticity in S-matrix theory," Commun. Math. Phys. **14** (1969), 15-55 doi:10.1007/BF01645454

[19] D. Atkinson, "The application of fixed-point theorems to S-matrix theory: problems of existence and uniqueness D. Atkinson," Lect. Theor. Phys. **11A** (1969) 523. [**?**]

[20] S. Mandelstam, "Determination of the pion - nucleon scattering amplitude from dispersion relations and unitarity. General theory," Phys. Rev. **112** (1958) 1344. doi:10.1103/PhysRev.112.1344

[21] S. Aks, "Proof that scattering implies production in quantum field theory," Journal of Mathematical Physics 6.4 (1965): 516-532.

[22] A. Dragt, "Amount of four-particle production required in S-matrix theory," Physical Review 156.5 (1967): 1588.

[23] Martin, A., 1962. Constraints Imposed on Mandelstam Representation by Unitarity. Physical Review Letters, 9(9), p.410.

[24] D. Atkinson, "A proof of the existence of functions that satisfy exactly both crossing and unitarity: I. Neutral pion-pion scattering. No subtractions.," Nucl. Phys. B **7** (1968), 375-408 [erratum: Nucl. Phys. B **15** (1970), 331-331] doi:10.1016/0550-3213(70)90120-3

[25] A. Martin and S. M. Roy, "Lower bound on inelasticity in pion-pion scattering," Phys. Rev. D **96** (2017) no.11, 114014 doi:10.1103/PhysRevD.96.114014 [arXiv:1710.07140 [hep-ph]].

[26] A. Martin, "Inability Of Field Theory To Exploit The Full Unitarity Condition," No. CERN-TH-727. CM-P00057359, 1966.

[27] M. Correia, A. Sever and A. Zhiboedov, "Probing multi-particle unitarity with the Landau equations," [arXiv:2111.12100 [hep-th]].

[28] S. Mizera, "Crossing symmetry in the planar limit," Phys. Rev. D **104** (2021) no.4, 045003 doi:10.1103/PhysRevD.104.045003 [arXiv:2104.12776 [hep-th]].

[29] M. F. Paulos, J. Penedones, J. Toledo, B. C. van Rees and P. Vieira, "The S-matrix bootstrap. Part III: higher dimensional amplitudes," JHEP **12** (2019), 040 doi:10.1007/JHEP12(2019)040 [arXiv:1708.06765 [hep-th]].

[30] A. Guerrieri and A. Sever, "Rigorous bounds on the Analytic S-matrix," [arXiv:2106.10257 [hep-th]].

[31] Y. He and M. Kruczenski, "S-matrix bootstrap in 3+1 dimensions: regularization and dual convex problem," JHEP **08** (2021), 125 doi:10.1007/JHEP08(2021)125 [arXiv:2103.11484 [hep-th]].

[32] S. Caron-Huot, Z. Komargodski, A. Sever and A. Zhiboedov, "Strings from Massive Higher Spins: The Asymptotic Uniqueness of the Veneziano Amplitude," JHEP **10** (2017), 026 doi:10.1007/JHEP10(2017)026 [arXiv:1607.04253 [hep-th]].

[33] S. Caron-Huot, D. Mazac, L. Rastelli and D. Simmons-Duffin, "Sharp Boundaries for the Swampland," JHEP **07** (2021), 110 doi:10.1007/jhep07(2021)110 [arXiv:2102.08951 [hep-th]].

[34] X. O. Camanho, J. D. Edelstein, J. Maldacena and A. Zhiboedov, "Causality Constraints on Corrections to the Graviton Three-Point Coupling," JHEP **02** (2016), 020 doi:10.1007/JHEP02(2016)020 [arXiv:1407.5597 [hep-th]].

[35] M. Correia, A. Sever and A. Zhiboedov, "An Analytical Toolkit for the S-matrix Bootstrap," doi:10.1007/JHEP03(2021)013 [arXiv:2006.08221 [hep-th]].

[36] A. Dymarsky, "Can Froissart Bound Explain Hadron Cross-Sections at High Energies?," JHEP **07** (2015), 106 doi:10.1007/JHEP07(2015)106 [arXiv:1412.8642 [hep-ph]].

[37] A. Adams, N. Arkani-Hamed, S. Dubovsky, A. Nicolis and R. Rattazzi, "Causality, analyticity and an IR obstruction to UV completion," JHEP **10** (2006), 014 doi:10.1088/1126-6708/2006/10/014 [arXiv:hep-th/0602178 [hep-th]].

[38] B. Bellazzini, J. Elias Miró, R. Rattazzi, M. Riembau and F. Riva, "Positive moments for scattering amplitudes," Phys. Rev. D **104** (2021) no.3, 036006 doi:10.1103/PhysRevD.104.036006 [arXiv:2011.00037 [hep-th]].

[39] L. Y. Chiang, Y. t. Huang, W. Li, L. Rodina and H. C. Weng, "Into the EFThedron and UV constraints from IR consistency," [arXiv:2105.02862 [hep-th]].

[40] S. Caron-Huot and V. Van Duong, "Extremal Effective Field Theories," JHEP **05** (2021), 280 doi:10.1007/JHEP05(2021)280 [arXiv:2011.02957 [hep-th]].

[41] A. J. Tolley, Z. Y. Wang and S. Y. Zhou, "New positivity bounds from full crossing symmetry," JHEP **05** (2021), 255 doi:10.1007/JHEP05(2021)255 [arXiv:2011.02400 [hep-th]].

[42] N. Arkani-Hamed, T. C. Huang and Y. T. Huang, "The EFT-Hedron," JHEP **05** (2021), 259 doi:10.1007/JHEP05(2021)259 [arXiv:2012.15849 [hep-th]].

[43] Z. Bern, D. Kosmopoulos and A. Zhiboedov, "Gravitational effective field theory islands, low-spin dominance, and the four-graviton amplitude," J. Phys. A **54** (2021) no.34, 344002 doi:10.1088/1751-8121/ac0e51 [arXiv:2103.12728 [hep-th]].

[44] A. A. Arkhipov, "Comment on Andre Martin: 'Can one improve the Froissart Bound?'," [arXiv:0903.3564 [hep-ph]].

[45] S. Caron-Huot, "Analyticity in Spin in Conformal Theories," JHEP **09** (2017), 078 doi:10.1007/JHEP09(2017)078 [arXiv:1703.00278 [hep-th]].

[46] L. F. Alday, J. Henriksson and M. van Loon, "Taming the ϵ-expansion with large spin perturbation theory," JHEP **07** (2018), 131 doi:10.1007/JHEP07(2018)131 [arXiv:1712.02314 [hep-th]].

[47] P. Dorey, "Exact S matrices," [arXiv:hep-th/9810026 [hep-th]].

[48] A. L. Fitzpatrick, J. Kaplan, D. Poland and D. Simmons-Duffin, "The Analytic Bootstrap and AdS Superhorizon Locality," JHEP **12** (2013), 004 doi:10.1007/JHEP12(2013)004 [arXiv:1212.3616 [hep-th]].

[49] Z. Komargodski and A. Zhiboedov, "Convexity and Liberation at Large Spin," JHEP **11** (2013), 140 doi:10.1007/JHEP11(2013)140 [arXiv:1212.4103 [hep-th]].

[50] S. Caron-Huot, D. Mazac, L. Rastelli and D. Simmons-Duffin, "Dispersive CFT Sum Rules," JHEP **05** (2021), 243 doi:10.1007/JHEP05(2021)243 [arXiv:2008.04931 [hep-th]].

[51] D. Atkinson, "Introduction to the use of non-linear techniques in s-matrix theory," Acta Phys. Austriaca Suppl. **7** (1970), 32-70 doi:10.1007/978-3-7091-5835-7_2